

---

# **Formulación de Algoritmo para Reconocimiento de Patrones en Señales del Mercado de Valores en la Bolsa de Nueva York**

**Alejandro Esteban Rendón Diosa**

---



**Universidad  
Tecnológica  
de Pereira**

**Pereira, Risaralda**



---

# **Formulación de Algoritmo para Reconocimiento de Patrones en Señales del Mercado de Valores en la Bolsa de Nueva York**

**Alejandro Esteban Rendón Diosa**

---

Facultad de Ingenierías  
de la Universidad Tecnológica de  
Pereira

Presentado por  
Alejandro Esteban Rendón Diosa  
en Pereira, Risaralda

Colombia. Diciembre, 2017

Director: Ramiro Andrés Barrios Valencia

Fecha de la Sustentación: Diciembre. 2017

# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Formulación del Problema</b>	<b>3</b>
<b>3. Justificación</b>	<b>5</b>
<b>4. Objetivos</b>	<b>7</b>
4.1. Objetivo general . . . . .	7
4.2. Objetivos específicos . . . . .	7
<b>5. Marco Referencial</b>	<b>9</b>
5.1. Marco Histórico . . . . .	9
5.2. Marco Conceptual . . . . .	9
5.3. Marco Legal . . . . .	11
<b>6. Diseño Metodológico</b>	<b>13</b>
<b>7. Delimitación del Algoritmo</b>	<b>15</b>
7.1. Estrategia . . . . .	15
7.2. Conocimientos Base . . . . .	15
7.2.1. Media Móvil (Moving Average) . . . . .	15
7.2.2. Cross Validation . . . . .	17
7.2.3. Grid Search . . . . .	18
7.3. Señales . . . . .	18
7.3.1. EMA 89-200 - Exponential Moving Average . . . . .	19
7.3.2. MACD - Moving Average Convergence/Divergence . . . . .	20
7.3.3. Oscilador Estocástico (Stochastic Oscillator) . . . . .	21
7.3.4. Índice de Fuerza Relativa (Relative Strength Index, RSI) . . . . .	24
7.4. Métodos . . . . .	25
7.4.1. Regresión Logística . . . . .	25
7.4.2. Linear Discriminant Analysis . . . . .	26
7.4.3. Quadratic Discriminant Analysis . . . . .	26
7.4.4. Support Vector Machine . . . . .	27

<b>8. Implementación</b>	<b>29</b>
8.1. Pre-procesamiento	29
8.1.1. Obtención de Datos	29
8.1.2. Normalización de Datos	31
8.1.3. Entrenamiento	31
8.2. Técnicas	31
8.2.1. LR - Logistic Regression	31
8.2.2. LDA - Linear Discriminant Analysis	31
8.2.3. QDA - Quadratic discriminant Analysis	32
8.2.4. SVM	32
8.3. Entrenamiento de Modelos	32
8.3.1. Cross validation	32
8.3.2. Grid Search	32
8.4. Pruebas sobre los Modelos	33
<b>9. Resultados</b>	<b>35</b>
9.1. Datos Ordenados	35
9.1.1. Entrenamiento	35
9.1.2. Pruebas	36
9.2. Datos Desordenados	37
9.2.1. Entrenamiento	37
9.2.2. Pruebas	38
<b>10. Conclusiones</b>	<b>41</b>
<b>A. Gráficas - Cross Validation</b>	<b>43</b>
<b>B. Parámetros en Funciones Sklearn</b>	<b>53</b>
B.1. Logistic Regression	53
B.2. Linear Discriminant Analysis	53
B.3. Quadratic Discriminant Analysis	53
B.4. SVM	53

# Índice de figuras

1.1. Influencia del precio de las acciones: Acciones, flujo de información y reacciones. [9] . . . . .	1
7.1. Ejemplo SMA 10, GOOGL . . . . .	17
7.2. Ejemplo EMA 89, GOOGL . . . . .	18
7.3. Ejemplo k-fold Cross Validation . . . . .	19
7.4. EMA 89 200, GOOGL . . . . .	20
7.5. MACD, GOOGL . . . . .	22
7.6. Ejemplo Oscilador Estocástico, GOOGL . . . . .	23
7.7. Ejemplo Índice de Fuerza Relativa, GOOGL . . . . .	24
7.8. Representación Gráfica: Regresión Logística, Iris Dataset . . . . .	25
7.9. Representación Gráfica: Linear Discriminant Analysis con dos clases . . . . .	26
7.10. Quadratic Discriminant Analysis con dos clases . . . . .	27
7.11. Representación Gráfica: Kernels SVM, Iris Dataset . . . . .	28
8.1. Preprocesamiento de datos [9] . . . . .	29
9.1. Resultados Grid Search . . . . .	35
9.2. Resultados Grid Search, Pruebas . . . . .	37
9.3. Resultados Grid Search - Datos desordenados . . . . .	38
9.4. Resultados Grid Search, Pruebas . . . . .	39
A.1. EMA 89 200, Logistic Regression . . . . .	43
A.2. EMA 89 200, LDA . . . . .	44
A.3. EMA 89 200, QDA . . . . .	44
A.4. EMA 89 200, SVM . . . . .	45
A.5. MACD, Logistic Regression . . . . .	45
A.6. MACD, LDA . . . . .	46
A.7. MACD, QDA . . . . .	46
A.8. MACD, SVM . . . . .	47
A.9. RSI, Logistic Regression . . . . .	47
A.10. RSI, LDA . . . . .	48
A.11. RSI, QDA . . . . .	48
A.12. RSI, SVM . . . . .	49

A.13.SO, Linear Regression . . . . .	49
A.14.SO, LDA . . . . .	50
A.15.SO, QDA . . . . .	50
A.16.SO, SVM . . . . .	51



# Listings

8.1. Formato de dataset . . . . .	30
8.2. Implementación de Logistic Regression en Sklearn . . . . .	31
8.3. Implementación de Linear Discriminant Analysis en Sklearn . . . . .	32
8.4. Implementación de Quadratic Discriminant Analysis en Sklearn . . . . .	32
8.5. Implementación de SVM en Sklearn . . . . .	32
8.6. Implementación de Cross Validation . . . . .	32
8.7. Parámetros Logistic Regression en Grid Search . . . . .	32
8.8. Parámetros LAD en Grid Search . . . . .	33
8.9. Parámetros QDA en Grid Search . . . . .	33
8.10. Parámetros SVM en Grid Search . . . . .	33
8.11. Ejemplo Implementación de Grid Search . . . . .	33
8.12. Ejemplo Prueba sobre Modelo resultante de Grid Search . . . . .	33



# Capítulo 1

## Introducción

Los operadores bursátiles (*también conocidos como traders*) intentan inventar una fórmula mágica para vencer al mercado. Muchos conceptos, encerrados en *Análisis Técnico* y *Análisis Fundamental*, son bastante conocidos, pero los ajustes que los hacen exitosos siguen siendo desconocidos.

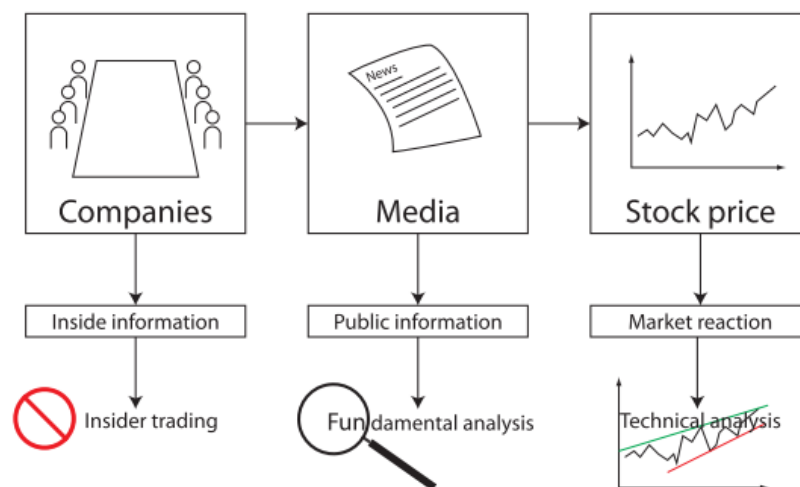


Figura 1.1: Influencia del precio de las acciones: Acciones, flujo de información y reacciones. [9]

En [Figura 1.1](#): Flujo de información de un *Índice bursátil*, desde las compañías hasta la interpretación de los movimientos.

En la actualidad se considera que la automatización de procesos es una prioridad en cualquier campo del conocimiento, ya que los costos están representados en el gasto de tiempo y en la presentación de resultados.

A partir de esta idea y conociendo de antemano que la Ingeniería de Sistemas tiene diversas aplicaciones en los distintos campos del conocimiento, se considera que es factible aplicar el enfoque hacia el campo de las Finanzas. El presente documento muestra la formulación de un proyecto que tiene como finalidad la implementación de una necesidad observada dentro de este campo; como lo es la de identificar patrones en señales de mercado de valores de la bolsa de Nueva York, aplicando técnicas de Inteligencia Artificial.

La Inteligencia Artificial brinda la ventaja de reducir costos, ya que muchas de las tareas que actualmente realizan los seres humanos podrían ser reemplazados para máquinas autosuficientes.

## Capítulo 2

# Formulación del Problema

En la actualidad se conocen algunos simuladores de Bolsa de Valores, como por ejemplo el de BVC [6]; los cuales son capaces de predecir con un gran índice de acierto los movimientos, pero con la desventaja de que son sistemas cerrados y accesibles solo a una población limitada.

La necesidad radica en que se quiere realizar un algoritmo que permita el reconocimiento de patrones en señales de mercado de valores en bolsa de Nueva York, ya que los métodos que son accesibles fácilmente son poco confiables y en la mayoría de casos no son acertados como es lo esperado.



# Capítulo 3

## Justificación

Por definición, las Finanzas son consideradas como el campo de la economía que se dedica de forma general al estudio de la obtención del dinero y su eficiente administración. El objetivo principal de las finanzas es el de ayudar a las personas naturales o jurídicas a tener un adecuado uso de su dinero, apoyándose en instrumentos financieros.

Con respecto a lo anterior, se puede decir que en un proceso financiero es de vital importancia el uso óptimo del tiempo entre la realización de operaciones. Por lo cual es muy factible la automatización de procesos por medio de la implementación de algoritmos que traten de predecir una salida altamente exitosa y con el mínimo tiempo de respuesta.





# Capítulo 4

## Objetivos

### 4.1. Objetivo general

Formular un algoritmo que permita identificar patrones en señales del mercado de valores de la bolsa de Nueva York, aplicando técnicas de Inteligencia Artificial.

### 4.2. Objetivos específicos

- Determinar 4 (cuatro) señales comunes en análisis técnico de índices bursátiles.
- Determinar técnicas de Inteligencia Artificial apropiadas para cada señal que haya sido identificada.
- Implementar las anteriores técnicas para cada señal.
- Evaluar eficacia de cada técnica implementada, utilizando comparaciones entre otros algoritmos planteados.



# Capítulo 5

## Marco Referencial

### 5.1. Marco Histórico

El origen de la Bolsa como institución data de finales del siglo XV en las ferias medievales de la Europa Occidental. En esas feria se inició la práctica de las transacciones de valores mobiliarios y títulos. El término “bolsa.”apareció en Brujas, Bélgica, concretamente en la familia de banqueros Van der Bursen, en cuyo palacio se organizó un mercado de títulos valores. En 1460 se creó la Bolsa de Amberes, que fue la primera institución bursátil en sentido moderno. Posteriormente, se creó la Bolsa de Londres en 1570, en 1595 la de Lyon, Francia y en 1792 la de Nueva York, siendo ésta la primera en el continente americano. Estas se consolidaron tras el auge de las sociedades anónimas.[2]

El índice estadounidense más antiguo existente hoy en día es el Dow Jones Industrial Average o simplemente conocido como Dow Jones. Éste mismo fue creado por Charles Henry Dow en conjunto con el Wall Street Journal para medir la actividad económica y financiera de los Estados Unidos de América a finales del siglo XIX. A principio estaba compuesto por tan solo de 12 compañías como General Electric, North America Company, American Tobacco Company entre otros, mientras que hoy en día está formado por 30 compañías.[3]

### 5.2. Marco Conceptual

Descripción de los conceptos, los cuales son fundamentales para el entendimiento de este proyecto

- *Índice Bursátil*: Instrumento de los Mercados Bursátiles para seguir la evolución del conjunto de las acciones cotizadas. La mayor parte de los índices se constituyen con una selección de acciones que pretenden representar a la totalidad.[2]

Los *índices bursátiles* son una referencia cada vez más importante para los gestores de cartera. Lo son también en la oferta de nuevos productos, sobre todo en depósitos

y fondos.[2]

■ *Análisis Técnico:*

Surgido a finales del siglo XIX, en Estados Unidos, a manos de Charles Henry Dow, este método prescinde del estudio de riesgo financiero propio a las sociedades y en su lugar estudia la acción del mercado, principalmente a través del uso de gráficas, con el propósito de predecir futuras tendencias en el precio.

Este estudio, que observa la cotización de la acción, el volumen bursátil y el interés abierto, utilizado principalmente en futuros y opciones, puede verse influenciado en gran medida por la propia perspectiva del analista.

La capacidad de determinar tendencias alcistas o bajistas complementa al análisis fundamental, haciendo que la suma de ambos tipos de análisis nos ayuden a comprender mejor hacia dónde se moverá el precio de las acciones [1].

■ *Análisis Fundamental:*

Introducido por Benjamin Graham y David Dodd, en 1934, este tipo de análisis trata de calcular el valor real de las acciones mediante los datos de los balances. Por supuesto, no realiza un análisis tan básico como la mera observación de los beneficios o pérdidas de la sociedad.

El análisis fundamental intenta calcular el valor real de un título mediante el análisis de balances y lo compara con el valor de mercado, para así descubrir si en un momento determinado el precio del mismo está infravalorado. Siempre teniendo en cuenta que los resultados de los balances y los precios se encuentran en constante cambio y que la cotización se puede ver influenciada también por otros factores entre los que se incluyen los políticos y los económicos [1].

- *OHLC:* La abreviatura establecida para *Open, High, Low, Close*, referente al precio de un instrumento/índice bursátil durante un tiempo dado.

- *Bolsa de Valores:* Las Bolsas de Valores se pueden definir como mercados organizados y especializados, en los que se realizan transacciones con títulos valores por medio de intermediarios autorizados, conocidos como Casas de Bolsa ó Puestos de Bolsa. Las Bolsas ofrecen al público y a sus miembros las facilidades, mecanismos e instrumentos técnicos que facilitan la negociación de títulos valores susceptibles de oferta pública, a precios determinados mediante subasta.

Dependiendo del momento en que un título ingresa al mercado, estas negociaciones se transarían en el mercado primario o en el mercado secundario. [2]

- *Mercado de Valores:* Es un subsistema dentro del sistema financiero y está compuesto por un conjunto de instrumentos o activos financieros, instituciones o intermediarios financieros cuya misión es contactar a compradores y vendedores en los mercados donde se negocian los diferentes instrumentos o activos financieros. [2]

- *Inteligencia Artificial*: Es la rama de las Ciencias de la Computación que estudia el software y hardware necesarios para simular el comportamiento y comprensión humanos.

El objetivo último de la IA es simular la inteligencia humana en una máquina creando robots que sean conscientes y con sentimientos reales, similares a los humanos. Uno de los problemas más difíciles es la simulación de la conciencia, cualidad humana que hace que nos demos cuenta de nuestra propia existencia. [4]

- *Hiperparámetro*: En el contexto del aprendizaje automático, los hiperparámetros son parámetros cuyos valores se establecen antes del comienzo del proceso de aprendizaje. Por el contrario, los valores de otros parámetros se derivan a través del entrenamiento.
- *Overffiting*: Es el efecto de sobreentrenar un algoritmo de aprendizaje con unos ciertos datos para los que se conoce el resultado deseado.

### 5.3. Marco Legal

A continuación se presentan un conjunto de leyes que regulan los Mercados de Valores según el estado Colombiano.

- **Ley 35 de 1993** Por la cual se dictan normas generales y se señalan en ellas los objetivos y criterios a los cuales debe sujetarse el Gobierno Nacional para regular las actividades financiera, bursátil y aseguradora y cualquier otra relacionada con el manejo, aprovechamiento e inversión de recursos captados del público y se dictan otras disposiciones en materia financiera y aseguradora.
- **Ley 1266 de 2008** Por medio de la cual se dictan disposiciones generales del Habeas Data y se regula el manejo de la información contenida en bases de datos personales, en especial la Financiera, Crediticia, Comercial, de Servicios y la proveniente de Terceros Países y se dictan otras disposiciones.
- **Ley 964 de 2005** Por la cual se dictan normas generales y se señalan en ellas los objetivos y criterios a los cuales debe sujetarse el Gobierno Nacional para regular las actividades de manejo, aprovechamiento e inversión de recursos captados del público que se efectúen mediante valores y se dictan otras disposiciones.
- **Ley 45 de 1990** Por la cual se expiden normas en materia de intermediación financiera, se regula la actividad aseguradora, se conceden unas facultades y se dictan otras disposiciones.
- **Ley 226 de 1995** Por la cual se desarrolla el artículo 60 de la Constitución Política en cuanto a la enajenación de la propiedad accionaria estatal, se toman medidas para su democratización y se dictan otras disposiciones.

- **Ley 795 de 2003** Por la cual se ajustan algunas normas del Estatuto Orgánico del Sistema Financiero y se dictan otras disposiciones.
- **Ley 546 de 1999** Por la cual se dictan normas en materia de vivienda, se señalan los objetivos y criterios generales a los cuales debe sujetarse el Gobierno Nacional para regular un sistema especializado para su financiación, se crean instrumentos de ahorro destinado a dicha financiación, se dictan medidas relacionadas con los impuestos y otros costos vinculados a la construcción y negociación de vivienda y se expiden otras disposiciones.
- **Ley 27 de 1990** Por la cual se dictan normas en relación con las bolsas de valores, el mercado público de valores, los depósitos centralizados de valores y las acciones con dividendo preferencial y sin derecho de voto.
- **Ley 1314 de 2009** Por la cual se regulan los principios y normas de contabilidad e información financiera y de aseguramiento de información aceptados en Colombia, se señalan las autoridades competentes, el procedimiento para su expedición y se determinan las entidades responsables de vigilar su cumplimiento
- **Ley 1328 de 2009** Por la cual se dictan normas en materia financiera, de seguros, del mercado de valores y otras disposiciones (Reforma Financiera).
- **Ley 510 de 1999** Por la cual se dictan disposiciones en relación con el sistema financiero y asegurador, el mercado público de valores, las Superintendencias Bancaria y de Valores y se conceden unas facultades.

# Capítulo 6

## Diseño Metodológico

La metodología propuesta para el desarrollo de la investigación se divide en cuatro aspectos fundamentales: *Delimitación del Algoritmo*, *Implementación*, *Análisis de Efectividad* y *Resultados*.

La etapa de *Delimitación del Algoritmo* consiste en describir la estrategia definida para el desarrollo de la investigación, así como la explicación de un grupo de componentes técnicos a tener en cuenta para la comprensión de dicho planteamiento.

A partir de las bases de conocimiento propuestas, se muestra una implementación de algunas *Técnicas de Inteligencia Artificial*, las cuales son utilizadas principalmente para Clasificación de características en conjuntos de datos.

Luego de tener una implementación óptima que logre interpretar nuestra solución planteada, se procede a realizar un *Análisis de Efectividad* para poder comparar los métodos a través de métricas y diferentes criterios de decisión. Posteriormente, se procede a sustentar los *Resultados* obtenidos del paso anterior.





# Capítulo 7

## Delimitación del Algoritmo

### 7.1. Estrategia

La manera en la que se decide avanzar en la sustentación de la hipótesis planteada se resumen en:

- Definición de conceptos técnicos requeridos para la comprensión de las respectivas implementaciones.
- Selección y descripción de señales comunes en análisis técnico de índices bursátiles ([Capítulo 4](#)).
- Selección y descripción de técnicas de Inteligencia Artificial apropiadas ([Capítulo 4](#))
- Implementar cada técnica sobre cada una de las señales.
- A partir de una métrica o criterio, decidir cuál técnica es la más efectiva para cada señal.

### 7.2. Conocimientos Base

En el [Capítulo 5](#) se vieron algunos fundamentos teóricos necesarios para el entendimiento básico de la presente investigación. Para complementar dichas bases se muestra a continuación una serie de fundamentos técnicos, los cuales serán utilizados en implementaciones de capítulos siguientes.

#### 7.2.1. Media Móvil (Moving Average)

El indicador técnico “Media Móvil” (*Moving Average*, *MA*) muestra el valor medio del precio de un instrumento durante un determinado período de tiempo. Cuando se calcula *Moving Average*, se promedia matemáticamente el precio del instrumento durante un período de tiempo dado. En función del cambio del precio, su valor medio va aumentando

o disminuyendo.

Tipos de medias móviles:

- *Simple Moving Average (SMA)* – Media Móvil Simple
- *Exponential Moving Average (EMA)* – Media Móvil Exponencial
- *Smoothed Moving Average (SMMA)* – Media Móvil Suavizada
- *Linear Weighted Moving Average (LWMA)* – Media Móvil Ponderada Lineal

**Nota:** Para los casos de estudio en los siguientes capítulos, solo es requerido el conocimiento sobre *MA* y *EMA*

### Media Móvil Simple (Simple Media Average, SMA)

La media móvil simple o aritmética se calcula mediante la suma de los precios de cierre de un instrumento durante un determinado número de períodos simples, dividiendo a continuación la suma por el número de estos períodos.

$$SMA = \frac{SUM(close(i), N)}{N} \quad (7.1)$$

Donde:

*SUM* - Suma

*Close(i)* - Precio de cierre del período actual

*N* - Número de periodos calculados

En la [Figura 7.1](#): Media Móvil Simple para un Periodo de 10 días para el índice GOOGL.

### Media Móvil Exponencial (Exponential Media Average, EMA)

La media móvil suavizada exponencialmente se calcula mediante la adición de una cierta parte del precio de cierre actual al valor anterior de la media móvil. A la hora de usar las medias móviles exponenciales, el mayor valor tienen los últimos precios de cierre.

$$EMA = (Close(i) * P) + (EMA(i - 1) * (1 - P)) \quad (7.2)$$

Donde:

*Close(i)* - Precio de cierre del período actual

*EMA(i - 1)* - valor de la media móvil para el período anterior

*P* - parte de uso del valor de precios.

En la [Figura 7.2](#): Media Móvil Exponencial para un Periodo de 89 días para el índice GOOGL.

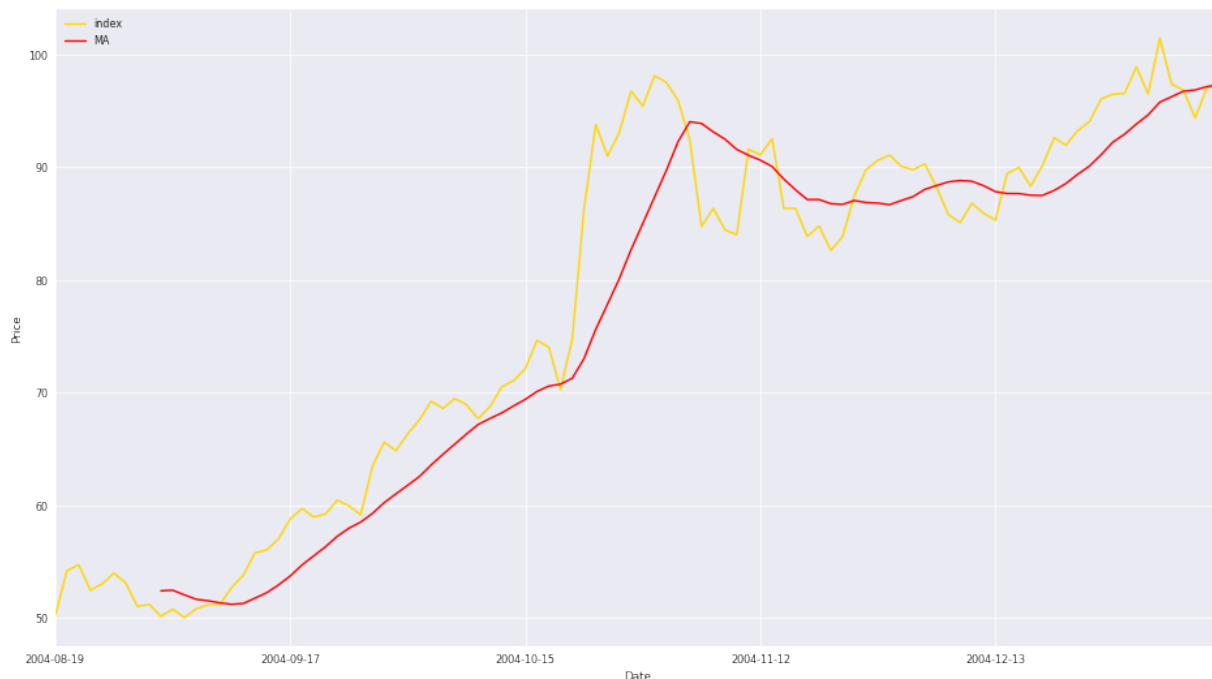


Figura 7.1: Ejemplo SMA 10, GOOGL

### 7.2.2. Cross Validation

Es un método estadístico de evaluación y comparación de algoritmos de aprendizaje, el cual divide el set de datos en dos partes: uno utilizado para el aprendizaje o entrenamiento de un modelo y el otro para validar el modelo. En un *Cross Validation* típico, los datos de entrenamiento y los datos de pruebas deben cruzarse en rondas consecutivas de modo que cada grupo de datos tiene la posibilidad de ser validado con el resto. La forma básica de *Cross Validation* es *K-fold Cross Validation* [10]

#### k-fold Cross validation

En *k-fold cross-validation* primero los datos son particionados en  $k$  segmentos o *folds* de igual tamaño o cercanos. Seguidamente  $k$  iteraciones de entrenamiento y de validación se realizan de tal manera que dentro de cada iteración un segmento diferente de los datos es separado para la validación mientras el resto de los  $k$  segmentos son usados para el entrenamiento. [10]

En Figura 7.3: Muestra un ejemplo con  $k = 5$ . La sección más oscura es utilizada para la validación del modelo, mientras que las cuatro secciones restantes son utilizadas para el entrenamiento del modelo.

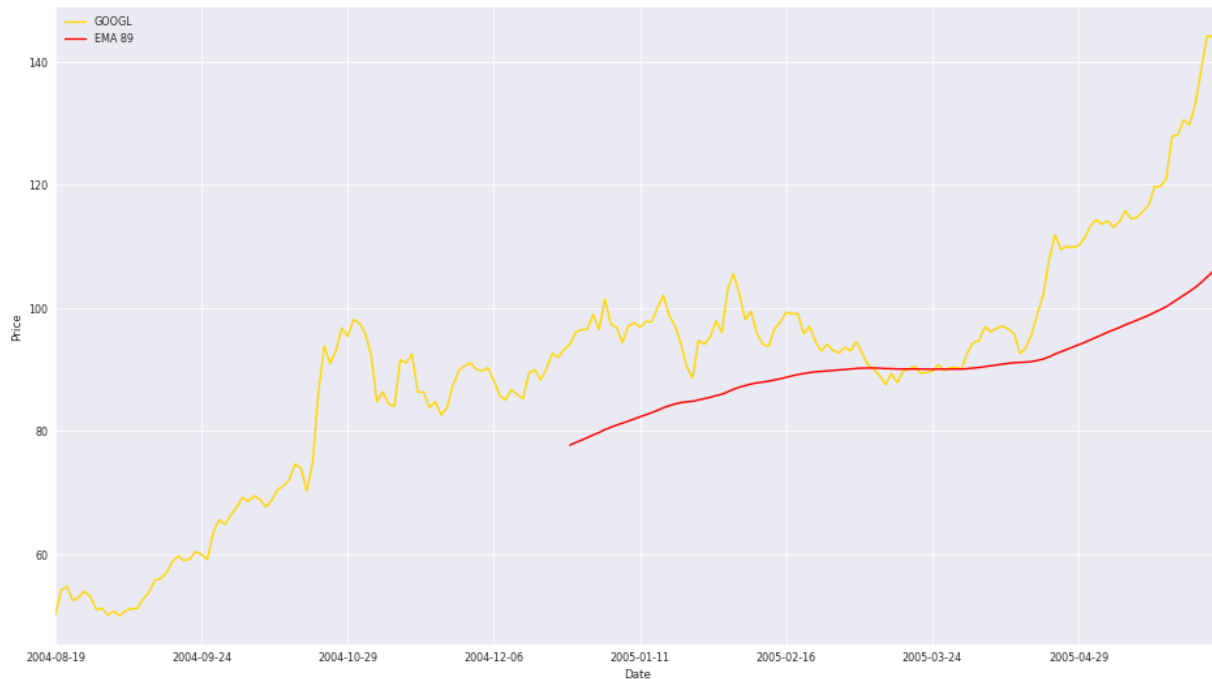


Figura 7.2: Ejemplo EMA 89, GOOGL

### 7.2.3. Grid Search

Es la forma tradicional de realizar la optimización de hiperparámetros, simplemente una búsqueda exhaustiva a través de un subconjunto especificado manualmente del espacio de un algoritmo de aprendizaje.

Debe guiarse por una métrica de rendimiento, generalmente medida por validación cruzada en el conjunto de entrenamiento o evaluación en un conjunto de validación extendido.

## 7.3. Señales

Según el [Capítulo 4](#), el primer paso crucial en la investigación es determinar 4 señales comunes en análisis técnico de índices bursátiles.

Basado en referencia del asesor de la investigación y considerando 4 de las señales más significativas para traders, la elección es la siguiente:

- EMA 89-200 - Exponential Moving Average, Periodos de 89 y 200 días
- MACD - Moving Average Convergence/Divergence
- RSI - Relative Strength Index
- SO - Stochastic Oscillator

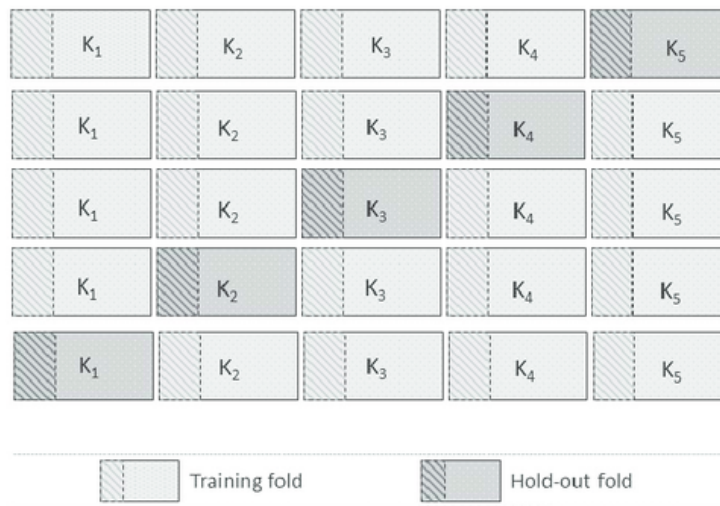


Figura 7.3: Ejemplo k-fold Cross Validation

En el desarrollo del presente capítulo se abordarán cada una de las señales anteriores, para contextualizar su importancia dentro del análisis técnico.

### 7.3.1. EMA 89-200 - Exponential Moving Average

La señal EMA 89-200 está constituida por un par de *EMAs* (Sección 7.2.1) con diferentes periodos. La función principal de esta señal es la de confirmar la tendencia.

En general, EMAs de 80 y 200 días son usadas como señales de trading de largos periodos.

#### Interpretación

- EMA 200 por debajo de EMA 89: la tendencia es alcista, los precios tienden a subir y es estable
- EMA 89 por debajo de EMA 200: la tendencia es bajista, es decir, los precios tienden a bajar y es bastante marcada
- Cruce pasando por encima de 89: Posiblemente hay un cambio de tendencia (de alcista a bajista, los precios no van a subir a bajar, comienzan a bajar)
- Cruce pasando por encima de 200: Posiblemente hay un cambio de tendencia (de bajista a alcista)

En la figura [Figura 7.4](#): Media Móvil Exponencial para Periodos de 89 y 200 días para el índice GOOGL.

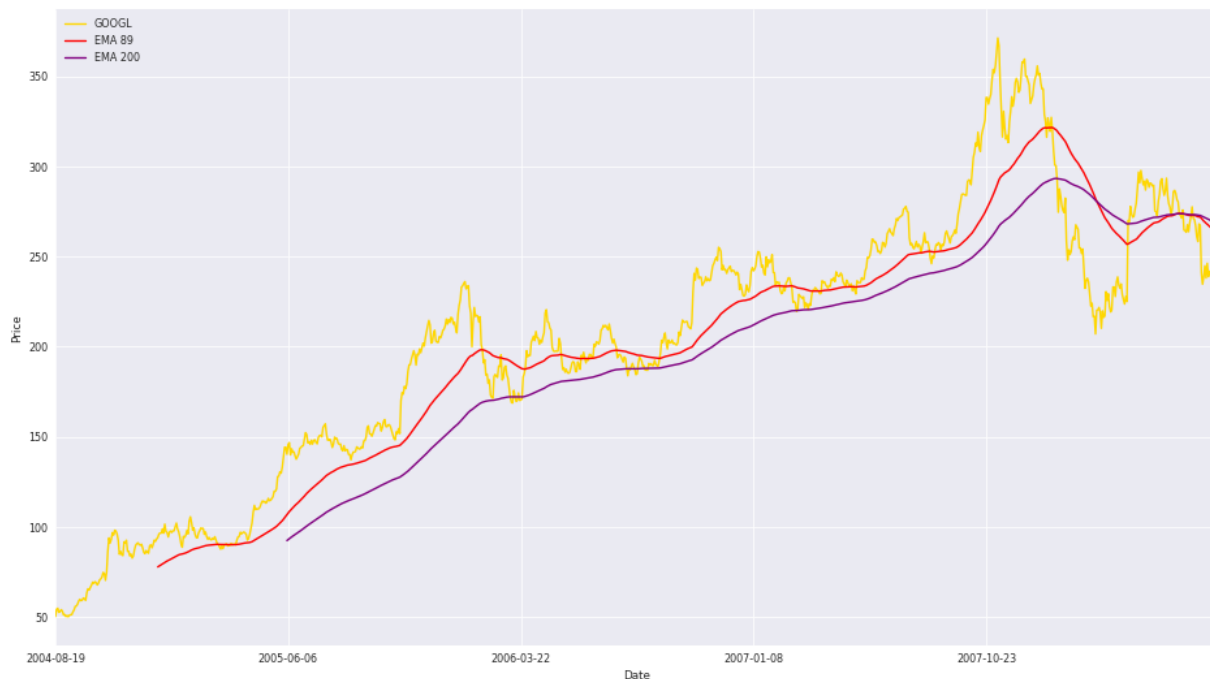


Figura 7.4: EMA 89 200, GOOGL

### 7.3.2. MACD - Moving Average Convergence/Divergence

El indicador técnico *Convergencia/Divergencia de Medias Móviles* es un indicador dinámico que sigue la tendencia. Muestra la correlación de dos medias móviles del precio.

El indicador técnico *MACD* se construye como la diferencia de dos media móviles exponenciales (*EMA*) con períodos de 12 y 26. Para definir con precisión los momentos oportunos de compra y venta, en el gráfico del MACD se dibuja una media móvil de 9 períodos, conocida como “*línea de señal*”. El MACD es más eficaz cuando el mercado se oscila con una gran amplitud dentro de la horquilla de trading.

Las señales más utilizados del MACD son las siguientes:

- Cruzamiento
- Estado de sobrecompra/sobreventa
- Divergencia

#### Cruzamiento

Cuando el Moving Average Convergence/Divergence cae por debajo de la línea de señal, hay que vender, y si sube por encima, hay que comprar. Cuando el MACD cruza la línea cero arriba/abajo, esto también puede servir de señales de compra/venta, respectivamente.

### Estado de sobrecompra/sobreventa

El Moving Average Convergence/Divergence es muy útil también como un indicador de sobrecompra/sobreventa. Cuando la media móvil corta sube sustancialmente por encima de la media móvil larga (es decir, el MACD va creciendo), eso significa que el precio del instrumento en cuestión seguramente está cotizado por encima de su valor real, y pronto volverá a un nivel más real.

### Divergencia

Cuando entre el MACD y el precio se forma una divergencia, significa la posibilidad de pronta finalización de la tendencia actual. Una divergencia alcista surge cuando el MACD alcanza nuevos máximos, y el precio no logra alcanzarlo. Una divergencia bajista surge cuando el indicador alcanza nuevos mínimos, y el precio no lo consigue. Los ambos tipos de divergencia tienen más valor si se forman en las áreas de sobrecompra/sobreventa.

### Fórmula

El indicador técnico Moving Average Convergence/Divergence se determina restando la media móvil exponencial de 26 períodos de la media móvil exponencial de 12 períodos. Luego por encima del gráfico del MACD se puntea una media móvil simple de 9 períodos (la línea de señal).

$$SIGNAL = SMA(MACD, 9) \quad (7.3)$$

$$MACD = EMA(CLOSE, 12) - EMA(CLOSE, 26) \quad (7.4)$$

Donde:

*EMA* – media móvil exponencial

*SMA* – media móvil simple

*SIGNAL* – línea de señal del indicador

En la figura [Figura 7.5](#): MACD y Señal EMA co Periodo de 9 días para el índice GOOGL.

### 7.3.3. Oscilador Estocástico (Stochastic Oscillator)

Compara el precio de cierre actual con el rango de precios durante un determinado período de tiempo. El indicador se compone de dos líneas. La línea principal se llama %K. La segunda línea %D es la media móvil de la línea %K. Normalmente la %K se muestra como una línea continua, y la %D – punteada. Existen tres modos más populares de interpretar el Oscilador estocástico:

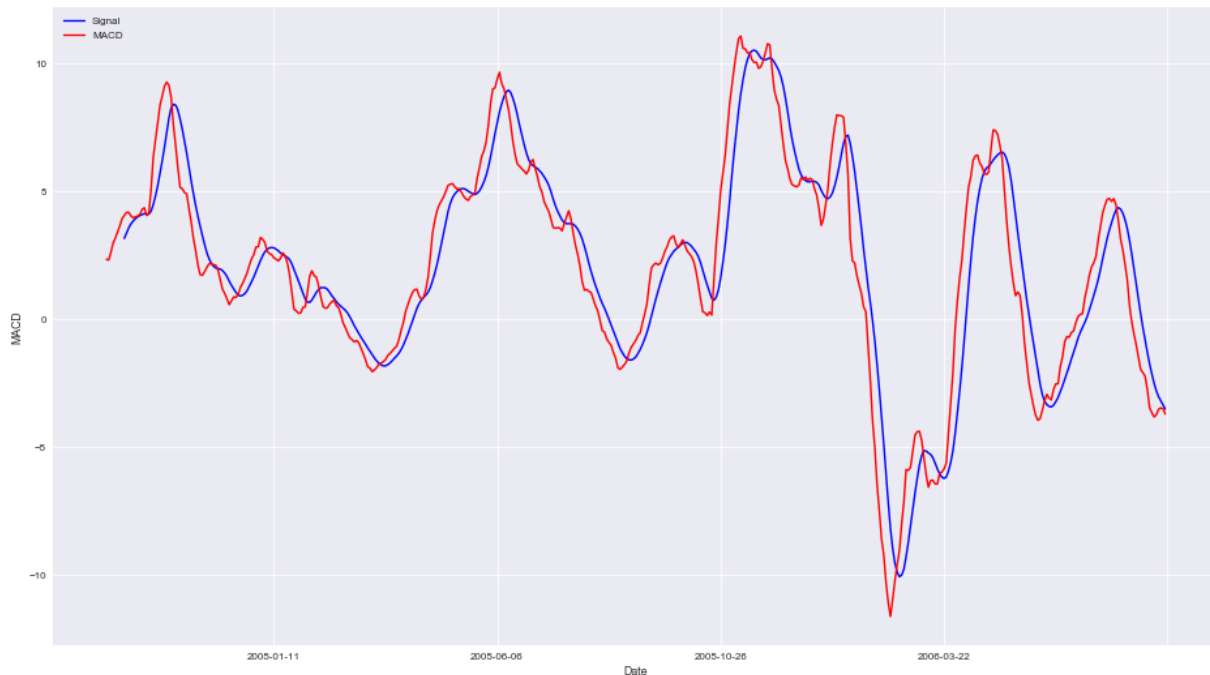


Figura 7.5: MACD, GOOGLE

- Compre cuando el oscilador ( $%K$  o  $%D$ ) primero baja por debajo de un determinado nivel (normalmente 20), y luego sube por encima de él. Venda cuando el oscilador primero se sube por encima de un determinado nivel (normalmente 80), y luego baja por debajo de él.
- Compre si la línea  $%K$  sube por encima de la  $%D$ . Venda si la línea  $%K$  baja por debajo de la  $%D$ .
- Vigile las desviaciones. Por ejemplo: los precios forman una nueva serie de los máximos, y el Stochastic Oscillator no logra subir por encima de sus máximos anteriores.

## Fórmula

Para calcular el Oscilador Estocástico se utilizan cuatro variables:

- Períodos  $%K$ . Es el número de períodos singulares que se utilizan para el cálculo del Oscilador Estocástico. Períodos de ralentización  $%K$ . Este valor determina el grado del suavizado interno de la línea  $%K$ . El valor 1 da el oscilador estocástico rápido, y el valor 3 – lento.
- Períodos  $%D$ . Es el número de períodos singulares que se utilizan para el cálculo de la media móvil de la línea  $%K$ .



- Método %D. Es el método de suavización (exponencial, simple, suavizado o ponderado) que se utiliza para el cálculo de %D.

Fórmula para calcular %K:

$$\%K = \frac{Close - MIN(LOW(\%K))}{MAX(HIGH(\%K)) - MIN(LOW(\%K))} * 100 \quad (7.5)$$

Aquí:

*Close* – precio de cierre de hoy;

*MIN(LOW(%K))* – el menor de los mínimos para el número de períodos %K;

*MAX(HIGH(%K))* – el mayor de los máximos para el número de períodos %K.

La media móvil de la %D se calcula mediante la siguiente fórmula:

$$\%D = SMA(\%K, N) \quad (7.6)$$

Aquí:

*N* – período de suavización;

*SMA* – media móvil simple.

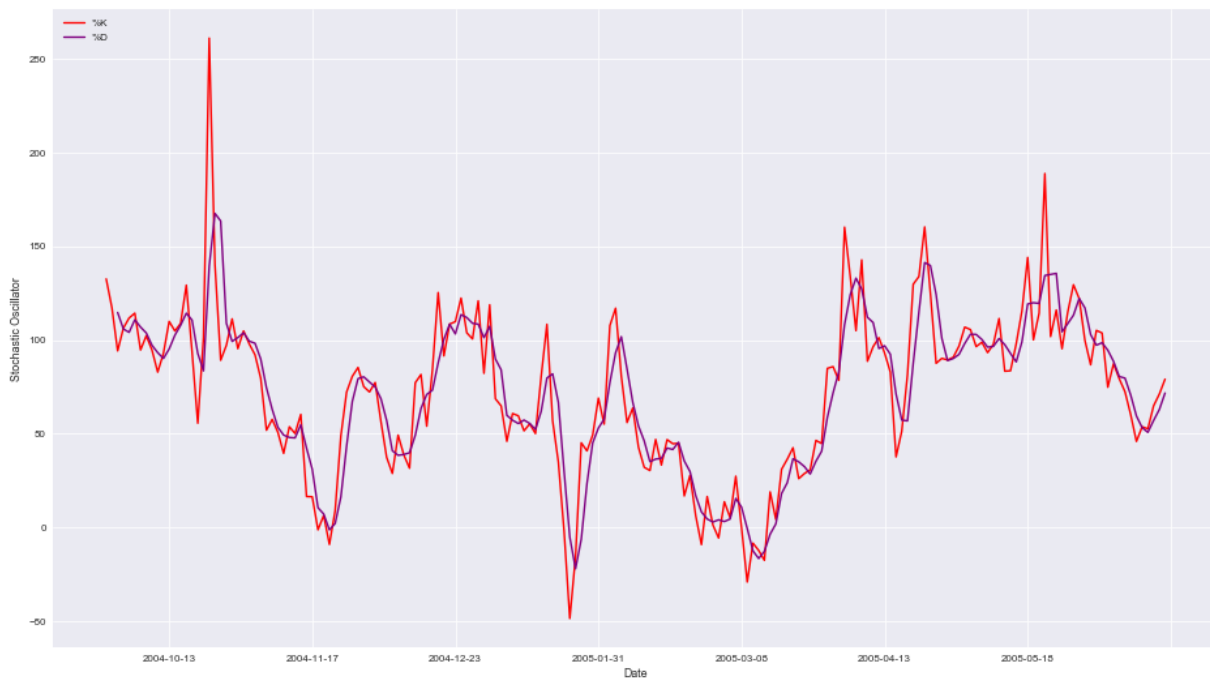


Figura 7.6: Ejemplo Oscilador Estocástico, GOOGL

En la figura [Figura 7.6](#): Periodos %K y %D de GOOGL.

### 7.3.4. Índice de Fuerza Relativa (Relative Strength Index, RSI)

Oscilador que sigue el precio y se oscila en el rango de 0 a 100. Uno de los métodos más populares de análisis del indicador Relative Strength Index consiste en la búsqueda de divergencias cuando el precio forma un nuevo máximo y el RSI no consigue superar el nivel de su máximo anterior. La divergencia de este tipo indica en una posible vuelta de precios. Si luego el indicador da la vuelta hacia abajo y cae por debajo de su mínimo, entonces se dice que finaliza una “oscilación fallida” (failure swing). Esta oscilación fallida es considerada como confirmación de un pronto giro de precios.

#### Fórmula

$$RSI = 100 - \left( \frac{100}{\left(1 + \frac{U}{D}\right)} \right) \quad (7.7)$$

Donde:

$U$  – valor medio de los cambios de precios positivos;

$D$  – valor medio de los cambios de precios negativos.

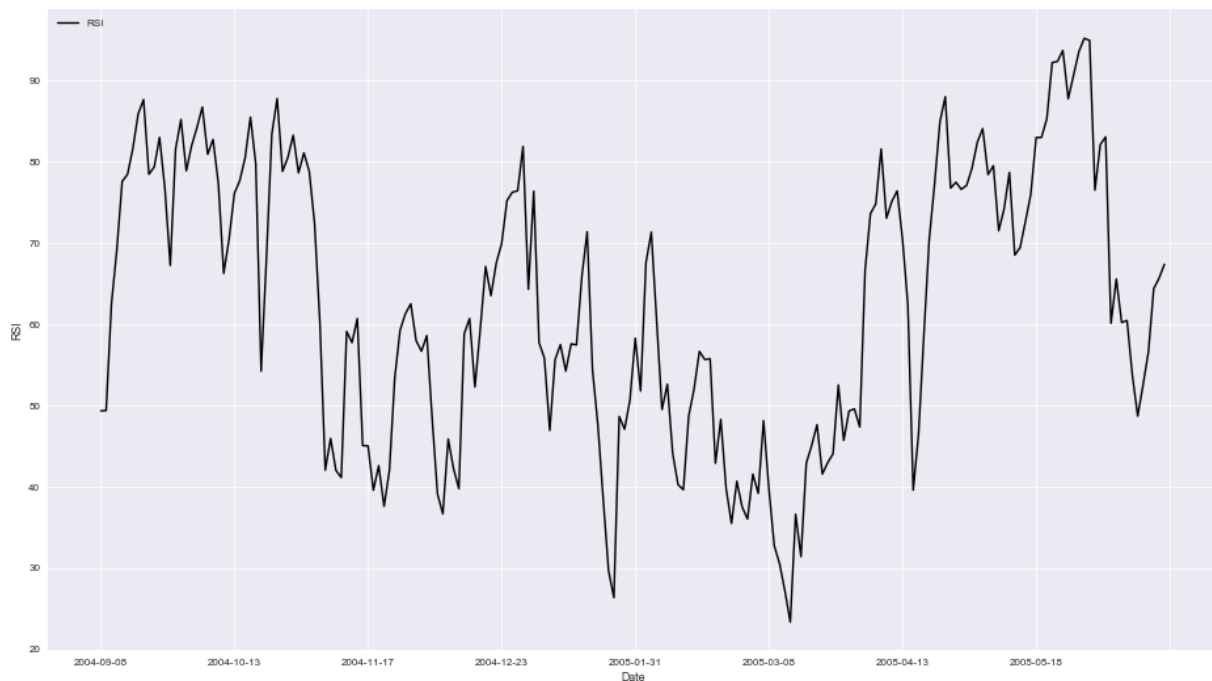


Figura 7.7: Ejemplo Índice de Fuerza Relativa, GOOGL

## 7.4. Métodos

Con base en investigaciones previas sobre el tema de *clasificación y detección de patrones* sobre índices bursátiles[12], se eligen las siguientes técnicas de *Machine Learning* para clasificación:

- Logistic Regression
- Linear Discriminant Analysis
- Quadratic Discriminant Analysis
- Support Vector Machine

*NOTA: En las secciones posteriores se analizará cada una de las técnicas listadas.*

### 7.4.1. Regresión Logística

Puede considerarse un caso especial del análisis de regresión en donde la variable dependiente es dicotómica y puede representarse de manera binaria.

Es por tanto, una técnica multivariante de dependencia ya que trata de estimar la probabilidad de que ocurra un suceso en función de la dependencia de otras variables.

Se trata de calcular la probabilidad en la que una de las opciones de la variable dicotómica dependiente sucederá en función de cómo puntúa en una serie de variables dependientes que pueden estar en diferentes escalas de medida [8]

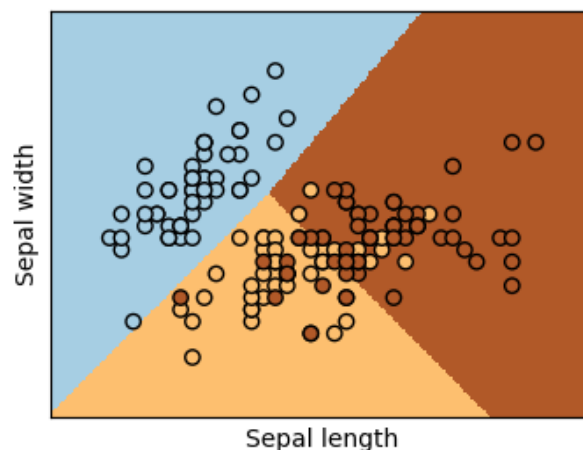


Figura 7.8: Representación Gráfica: Regresión Logística, Iris Dataset

En [Figura 7.8](#): Representación gráfica de una Regresión Logística con el dataset clásico *Iris*[5]

### 7.4.2. Linear Discriminant Analysis

Es usado en estadística, reconocimiento de patrones y Machine Learning para encontrar una combinación lineal las cuales caracterizan o separan dos o más clases de objetos o eventos. [13].

*LDA* está también estrechamente relacionado con el análisis de componente principal y el análisis factorial en que ambos buscan combinaciones lineales de variables que explican mejor los datos. *LDA* explícitamente intenta modelar la diferencia entre las clases de datos.

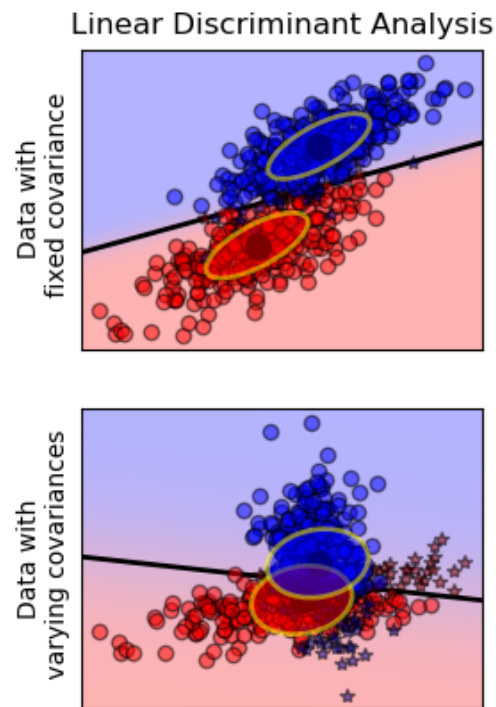


Figura 7.9: Representación Gráfica: Linear Discriminant Analysis con dos clases

En [Figura 7.9](#): Aplicación de Linear Discriminant Analysis para Clasificar en dos características.

### 7.4.3. Quadratic Discriminant Analysis

Está estrechamente relacionado con el Análisis Discriminante Lineal (LDA), donde se supone que las mediciones de cada clase se distribuyen normalmente. Sin embargo, a diferencia de LDA, en QDA no se supone que la covarianza de cada una de las clases sea idéntica.

Cuando la suposición de normalidad es verdadera, la mejor prueba posible para la hipótesis de que una medida dada es de una clase dada es la prueba de razón de verosimilitud.

## Quadratic Discriminant Analysis

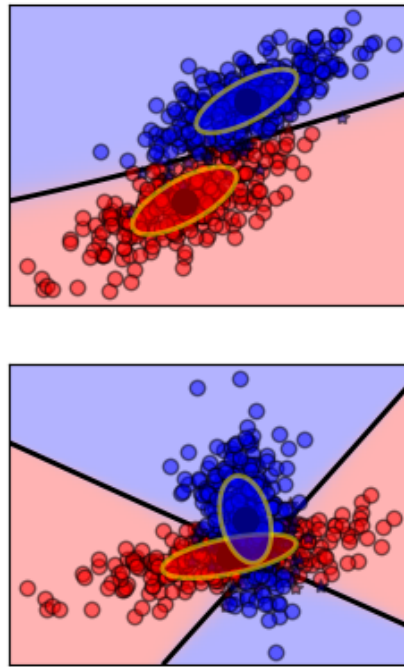


Figura 7.10: Quadratic Discriminant Analysis con dos clases

En [Figura 7.10](#): Aplicación de Quadratic Discriminant Analysis para Clasificar en dos características.

#### 7.4.4. Support Vector Machine

El método *Máquina de Soporte Vectorial* - *SVM* está propiamente relacionado con problemas de clasificación y regresión.

Dado un conjunto de ejemplos de entrenamiento (de muestras) podemos etiquetar las clases y entrenar una SVM para construir un modelo que prediga la clase de una nueva muestra. Intuitivamente, una SVM es un modelo que representa a los puntos de muestra en el espacio, separando las clases a 2 espacios lo más amplios posibles mediante un hiperplano de separación definido como el vector entre los 2 puntos, de las 2 clases, más cercanos al que se llama vector soporte. Cuando las nuevas muestras se ponen en correspondencia con dicho modelo, en función de los espacios a los que pertenezcan, pueden ser clasificadas a una o la otra clase.

En [Figura 7.11](#): Clasificación de conjunto de datos [5] utilizando diferentes Kernels con SVM.

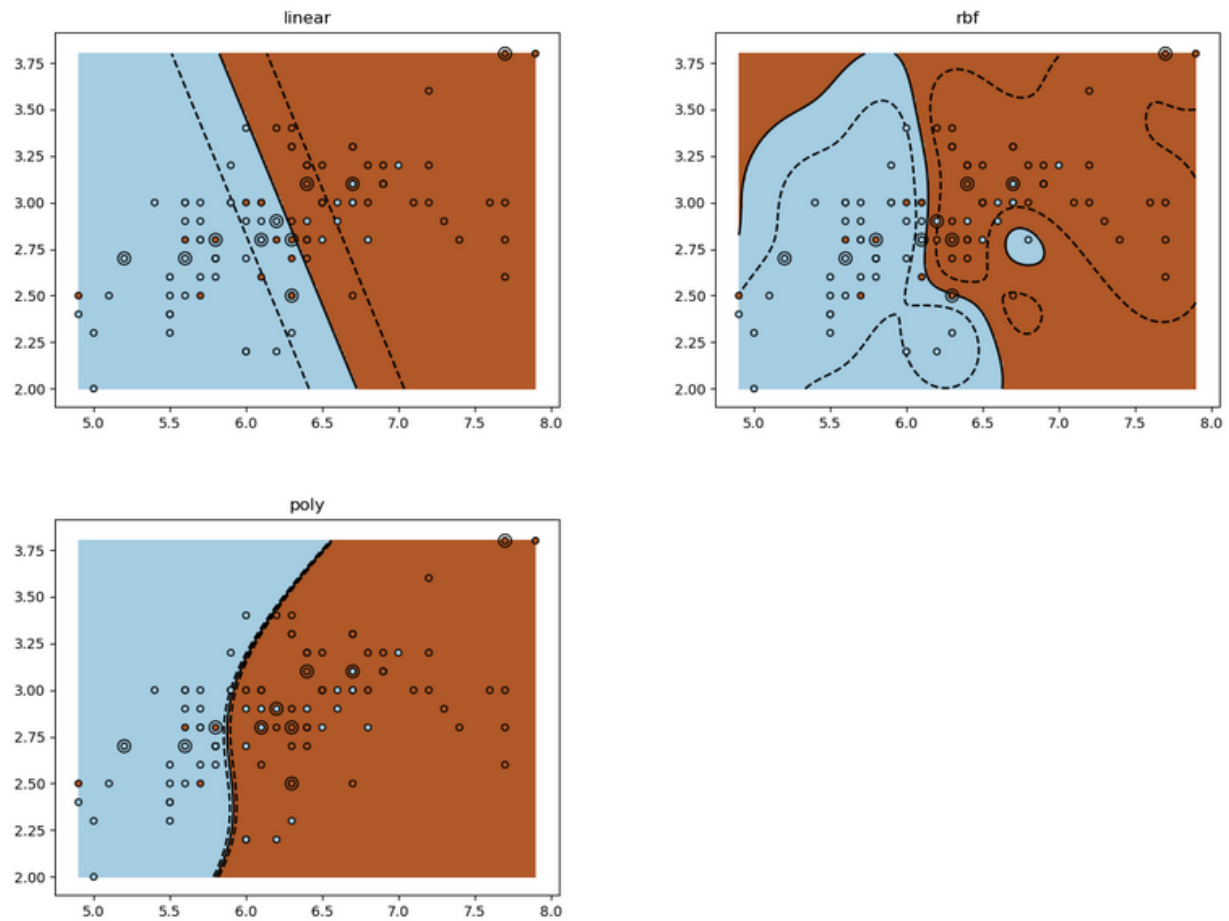


Figura 7.11: Representación Gráfica: Kernels SVM, Iris Dataset

# Capítulo 8

## Implementación

### 8.1. Pre-procesamiento

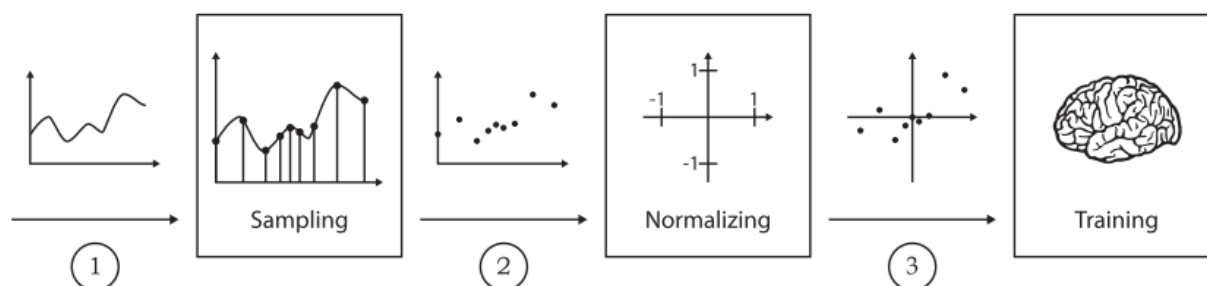


Figura 8.1: Preprocesamiento de datos [9]

En [Figura 8.1](#): Flujo de trabajo con grupos de datos antes del entrenamiento del modelo.

#### 8.1.1. Obtención de Datos

El primer paso a seguir es obtener un conjunto de datos para analizar.

En el caso de los índices bursátiles solo es requerida la información *OHLC*. Para el caso de estudio a abordar se requiere la información diaria de los índices.

Dicha información está disponible en las diferentes bolsas de valores. Para la investigación se eligieron algunos índices de la Bolsa de Nueva York (*FB* - *Facebook*, *GOOGL* - *Google* y *TWTR* - *Twitter*).

Existen múltiples maneras de obtener los conjuntos de datos ya que hay varios proveedores conectados a la bolsa y prestan el servicio de recolección y visualización de dichos datos, como por ejemplo las compañías Yahoo y Google; ambas tienen una división de Finanzas, la cual almacena y brinda herramientas al público sobre actualidad financiera.

Las dos alternativas más reconocidas para la obtención de los recursos son:

- Consultas sobre la *API* de cada compañía.
- Descarga de recurso directamente en un texto plano

### Formato de los datos

```

1 Date: {
2     'Open': '',
3     'High': '',
4     'Low': '',
5     'Close': '',
6     'Volume': ''
7 } ...

```

Listing 8.1: Formato de dataset

### Descripción de los Datos Seleccionados

- EMA 89 200
  - Tamaño:
    - Entrenamiento: 4405 días
    - Pruebas: 750 días
  - Campos: *OHLC, EMA\_89, EMA\_200, Label (EMA\_89, EMA\_200, Intersection)*
- MACD:
  - Tamaño:
    - Entrenamiento: 4853 días
    - Pruebas: 800 días
  - Campos: *OHLC, MACD, signal\_MACD, Label (Up, Down, Intersection)*
- RSI:
  - Tamaño:
    - Entrenamiento: 4913 días
    - Pruebas: 800 días
  - Campos: *OHLC, rsi, Label (30, 70, 0)*
- SO:
  - Tamaño:
    - Entrenamiento: 4954 días
    - Pruebas: 750 días
  - Campos: *OHLC, k, d, Label (K, D, Intersection)*



### 8.1.2. Normalización de Datos

En estadística y las aplicaciones de estadísticas, la normalización puede tener una gama de significados.

En los casos más sencillos, la normalización de índices significa ajustar los valores medidos en diferentes escalas respecto a una escala común, a menudo previo a un proceso de realizar promedios. En casos más complicados, la normalización puede referirse ajustes más sofisticados donde la intención es conseguir todas las distribuciones de probabilidad que se ajustan a los valores

En el caso de normalización de puntuaciones en valoración educativa, puede haber una intención para alinear distribuciones a una distribución normal. Una aproximación diferente a la normalización de distribuciones de probabilidad es normalización de cuantiles, donde los cuantiles de diferentes medidas son traídas al ajuste [11]

### 8.1.3. Entrenamiento

También llamada la etapa de aprendizaje, es el proceso en el que se detectan los patrones de un conjunto de datos, es decir, es el corazón del machine learning. Una vez identificados los patrones, se pueden hacer predicciones con nuevos datos que se incorporen al sistema.

## 8.2. Técnicas

Para las respectivas pruebas que sustentan los planteamientos de esta investigación, se muestra una implementación de las técnicas elegidas en Sklearn [7].

*Sklearn* es una librería de *Machine Learning* gratuita de software libre para el lenguaje de Programación Python. Es caracterizado por las implementaciones de varios algoritmos de Clasificación, Regresión y Clustering.

A continuación se expone cada una de las funciones para cada técnica seleccionada en (Sección 7.3), cada una con los parámetros por defecto según la implementación en Sklearn [7]:

### 8.2.1. LR - Logistic Regression

```
1 class sklearn.linear_model.LogisticRegression(penalty='l2', dual=False,
2         tol=0.0001, C=1.0, fit_intercept=True, intercept_scaling=1,
3         class_weight=None, random_state=None, solver='liblinear',
4         max_iter=100, multi_class='ovr', verbose=0, warm_start=False,
5         n_jobs=1)
```

Listing 8.2: Implementación de Logistic Regression en Sklearn

### 8.2.2. LDA - Linear Discriminant Analysis

```

1 class sklearn.discriminant_analysis.LinearDiscriminantAnalysis(solver='svd',
2     shrinkage=None, priors=None, n_components=None,
3     store_covariance=False, tol=0.0001

```

Listing 8.3: Implementación de Linear Discriminant Analysis en Sklearn

### 8.2.3. QDA - Quadratic discriminant Analysis

```

1 class sklearn.discriminant_analysis.QuadraticDiscriminantAnalysis(
2     priors=None, reg_param=0.0, store_covariance=False,
3     tol=0.0001, store_covariances=None)

```

Listing 8.4: Implementación de Quadratic Discriminant Analysis en Sklearn

### 8.2.4. SVM

```

1 class sklearn.svm.SVC(C=1.0, kernel='rbf', degree=3, gamma='auto',
2     coef0=0.0, shrinking=True, probability=False,
3     tol=0.001, cache_size=200, class_weight=None, verbose=False,
4     max_iter=-1, decision_function_shape='ovr', random_state=None)

```

Listing 8.5: Implementación de SVM en Sklearn

## 8.3. Entrenamiento de Modelos

### 8.3.1. Cross validation

```

1 train_sizes, train_scores, cv_scores = learning_curve(clf, X_train,
2     y_train.values.ravel(), scoring='accuracy',
3     train_sizes=train_sizes, n_jobs=2)

```

Listing 8.6: Implementación de Cross Validation

### 8.3.2. Grid Search

Cómo el objetivo de la técnica es encontrar el ajuste de parámetros para cada clasificador elegido. A continuación se lista el conjunto de parámetros y las opciones por cada uno (Ver parámetros según la documentación [Apéndice B](#)):

- Logistic Regression

```

1 {
2     'logit_reg__penalty': ['l1', 'l2'],
3     'logit_reg__C': [0.001, 0.01, 0.1, 1, 10, 100, 1000],
4     'logit_reg__tol': [0.00001, 0.0001, 0.001, 0.01, 0.1, 1],
5     'logit_reg__max_iter': [1, 10, 100]
6 }
7

```

Listing 8.7: Parámetros Logistic Regression en Grid Search

- LAD

```

1  {
2      'linearD__tol': [0.00001, 0.0001, 0.001, 0.01, 0.1, 1]
3  }
4

```

Listing 8.8: Parámetros LAD en Grid Search

#### ■ QDA

```

1  {
2      'QuadraticD__tol': [0.00001, 0.0001, 0.001, 0.01, 0.1, 1]
3  }
4

```

Listing 8.9: Parámetros QDA en Grid Search

#### ■ SVM

```

1  {
2      'SVM__kernel': ['linear', 'rbf', 'sigmoid', 'poly'],
3      'SVM__tol': [0.00001, 0.0001, 0.001, 0.01, 0.1, 1],
4      'SVM__gamma': [0.00001, 0.0001, 0.001, 0.01, 0.1, 1]
5  }
6

```

Listing 8.10: Parámetros SVM en Grid Search

La implementación básica de un Grid Search se muestra a continuación, dado un conjunto de parámetros como los del punto anterior:

```

1 method = Function()
2 pipe = Pipeline([( 'method', method)])
3
4 param_grid = {
5     '': [ '' ],
6     ...
7 }
8 grid = GridSearchCV(pipe, param_grid, scoring='accuracy', n_jobs=4)
9
10 gridLogistic.fit(X, y)

```

Listing 8.11: Ejemplo Implementación de Grid Search

## 8.4. Pruebas sobre los Modelos

Luego del entranamiento del modelo por medio de la técnica Grid Search, se procede a probar con este un grupo de datos constituidos únicamente por *OHLC*, de los cuales ya se sabe el *Label* pero que el modelo no conoció en la etapa de entrenamiento.

```

1 y_pred = gridLogistic.predict(X_test)
2 accuracy = accuracy_score(y_test, y_pred)

```

Listing 8.12: Ejemplo Prueba sobre Modelo resultante de Grid Search



# Capítulo 9

## Resultados

Con la implementación [Subsección 7.2.3](#) se obtuvieron los siguientes resultados (Medición por *'accuracy'*):

### 9.1. Datos Ordenados

#### 9.1.1. Entrenamiento

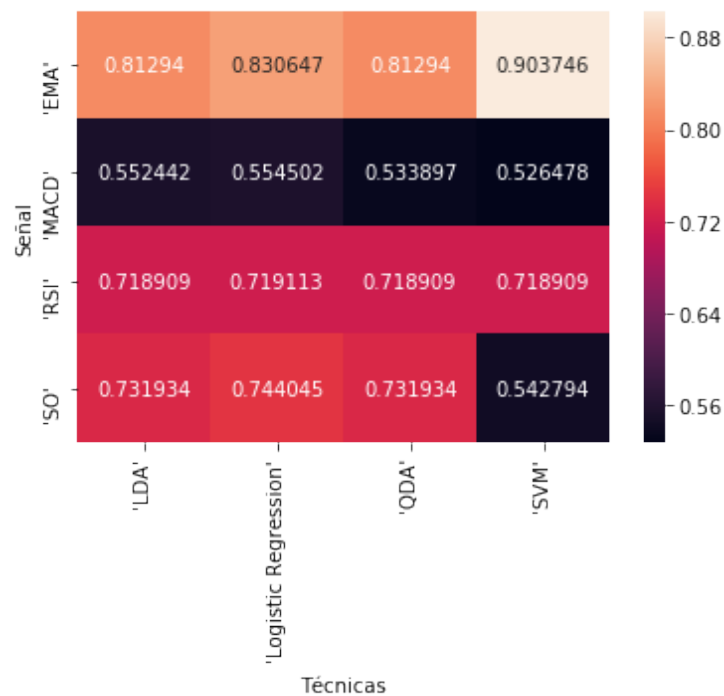


Figura 9.1: Resultados Grid Search

### Mejor resultado por Señal

- EMA 89 200:
  - Técnica: SVM
  - Accuracy: 0.903746
  - Parámetros: 'SVM\_gamma': 1, 'SVM\_kernel': 'rbf', 'SVM\_tol': 0.01
- MACD:
  - Técnica: Logistic Regression
  - Accuracy: 0.554502369668
  - Parámetros: 'logit\_reg\_C': 1000, 'logit\_reg\_max\_iter': 100, 'logit\_reg\_penalty': 'l1', 'logit\_reg\_tol': 1e-05
- RSI:
  - Técnica: Logistic Regression
  - Accuracy: 0.719112558518
  - Parámetros: 'logit\_reg\_C': 100, 'logit\_reg\_max\_iter': 10, 'logit\_reg\_penalty': 'l2', 'logit\_reg\_tol': 0.01
- SO:
  - Técnica: Logistic Regression
  - Accuracy: 0.744045215987
  - Parámetros: 'logit\_reg\_C': 10, 'logit\_reg\_max\_iter': 100, 'logit\_reg\_penalty': 'l1', 'logit\_reg\_tol': 0.0001

### 9.1.2. Pruebas

#### Mejor resultado por Señal

- EMA 89 200:
  - Accuracy: 0.569333
- MACD:
  - Accuracy: 0.5475
- RSI:
  - Accuracy: 0.70125
- SO:
  - Accuracy: 0.782667

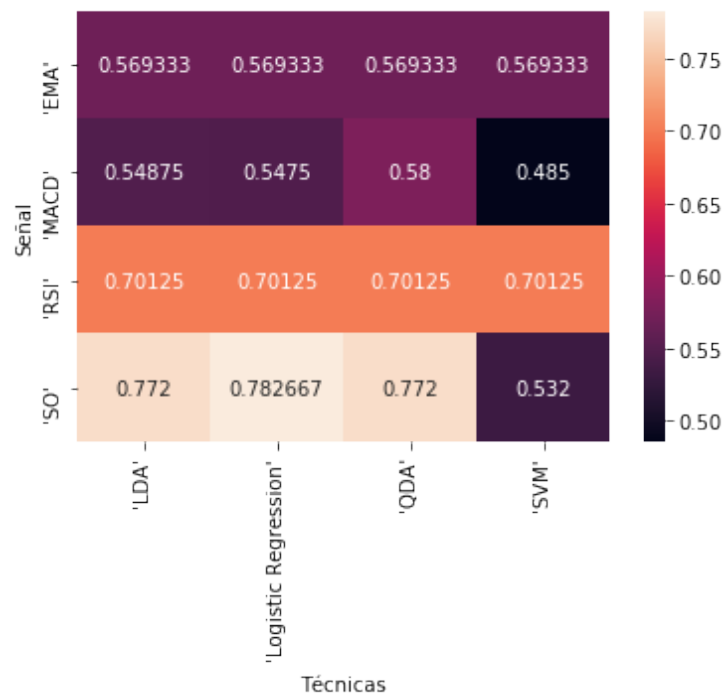


Figura 9.2: Resultados Grid Search, Pruebas

## 9.2. Datos Desordenados

### 9.2.1. Entrenamiento

Mejor resultado por Señal

- EMA 89 200:
  - Técnica: SVM
  - Accuracy: 0.862202043133
  - Parámetros: 'SVM\_\_gamma': 1, 'SVM\_\_kernel': 'rbf', 'SVM\_\_tol': 1e-05
- MACD:
  - Técnica: Logistic Regression
  - Accuracy: 0.551205439934
  - Parámetros: 'logit\_reg\_\_C': 10, 'logit\_reg\_\_max\_iter': 100, 'logit\_reg\_\_penalty': 'l1', 'logit\_reg\_\_tol': 1e-05
- RSI:
  - Técnica: Logistic Regression, LDA, QDA, SVM

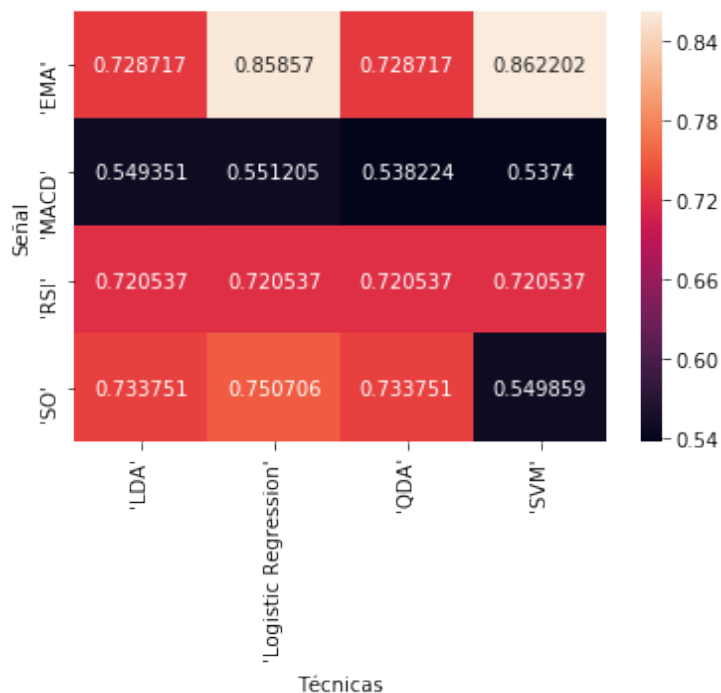


Figura 9.3: Resultados Grid Search - Datos desordenados

- Accuracy: 0.720537349888
- Parámetros: 'logit\_reg\_C': 0.001, 'logit\_reg\_max\_iter': 1, 'logit\_reg\_penalty': 'l1', 'logit\_reg\_tol': 1e-05, 'linearD\_tol': 0.1, 'linearD\_tol': 0.1, 'SVM\_gamma': 1e-05, 'SVM\_kernel': 'linear', 'SVM\_tol': 1e-05
- SO:
  - Técnica: LDA, QDA
  - Accuracy: 0.750706499798
  - Parámetros: 'linearD\_tol': 1e-05

### 9.2.2. Pruebas

#### Mejor resultado por Señal

- EMA 89 200:
  - Accuracy: 0.8586666666666667
- MACD:
  - Accuracy: 0.53375



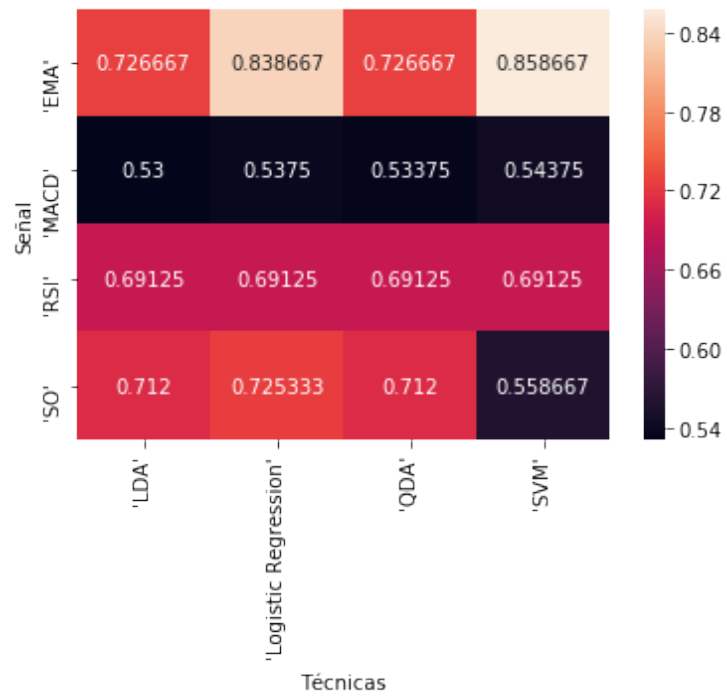


Figura 9.4: Resultados Grid Search, Pruebas

- RSI:
  - Accuracy: 0.69125
- SO:
  - Accuracy: 0.7253333333333334



# Capítulo 10

## Conclusiones

- Observando los resultados [Capítulo 9](#) se ve como el modelo del conjunto de datos ordenados por Fecha está *overfitting*, es decir, el *accuracy* tiene un buen puntaje en aquellas pruebas realizadas en el mismo entrenamiento, mientras que con la parte de datos desconocida para el modelo es bastante bajo.
- El puntaje tan bajo de *accuracy* es debido al tamaño del conjunto de datos elegido para las pruebas ya que el modelo no conoce algunas características importantes del índice bursátil
- En el caso de comparación entre datos ordenados y no ordenados por fecha, se tiene como resultado para nuestro caso de estudio que la aleatoriedad en el orden de los datos produce un mayor puntaje de *accuracy* en las pruebas.
- La técnica adecuada para cada Señal:
  - EMA 89 200: LDA y QDA
  - MACD: SVM
  - RSI: Logistic Regression, LDA, QDA y SVM
  - SO: Logistic Regression
- El algoritmo sugerido:
  1. Obtener Datos
  2. Análisis de datos
  3. Elección de señales
  4. Implementación de señales
  5. Elección de técnicas para clasificación
  6. Aplicar técnicas de clasificación a señales, utilizando *k-fold Cross Validation* dentro de un *Grid Search*, con el fin de obtener un modelo más justo y preciso

7. Por último, validar el modelo con un conjunto de datos diferente a los del entrenamiento en el punto 6

# Apéndice A

## Gráficas - Cross Validation

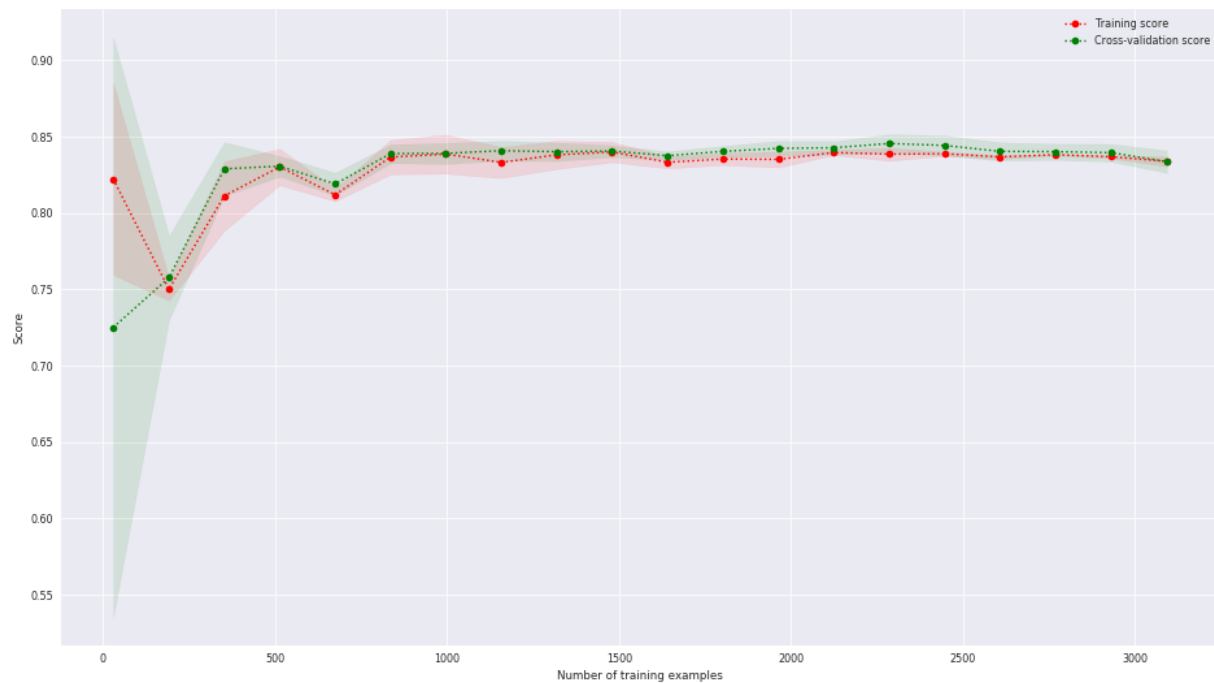


Figura A.1: EMA 89 200, Logistic Regression

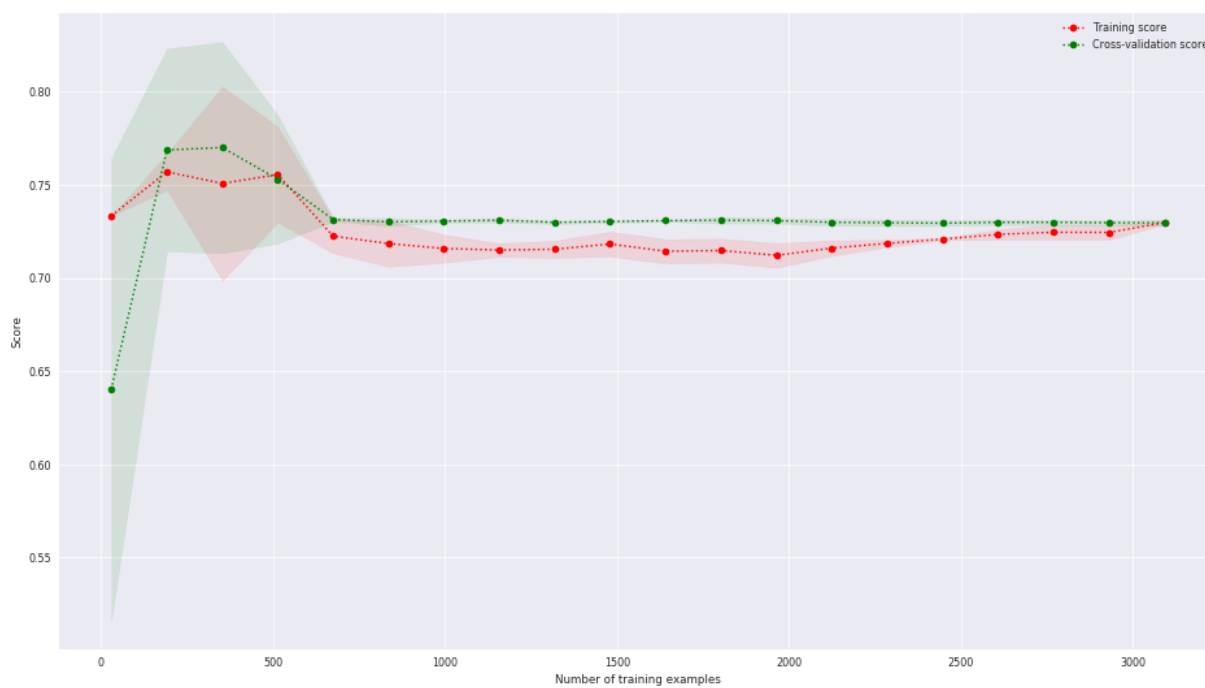


Figura A.2: EMA 89 200, LDA

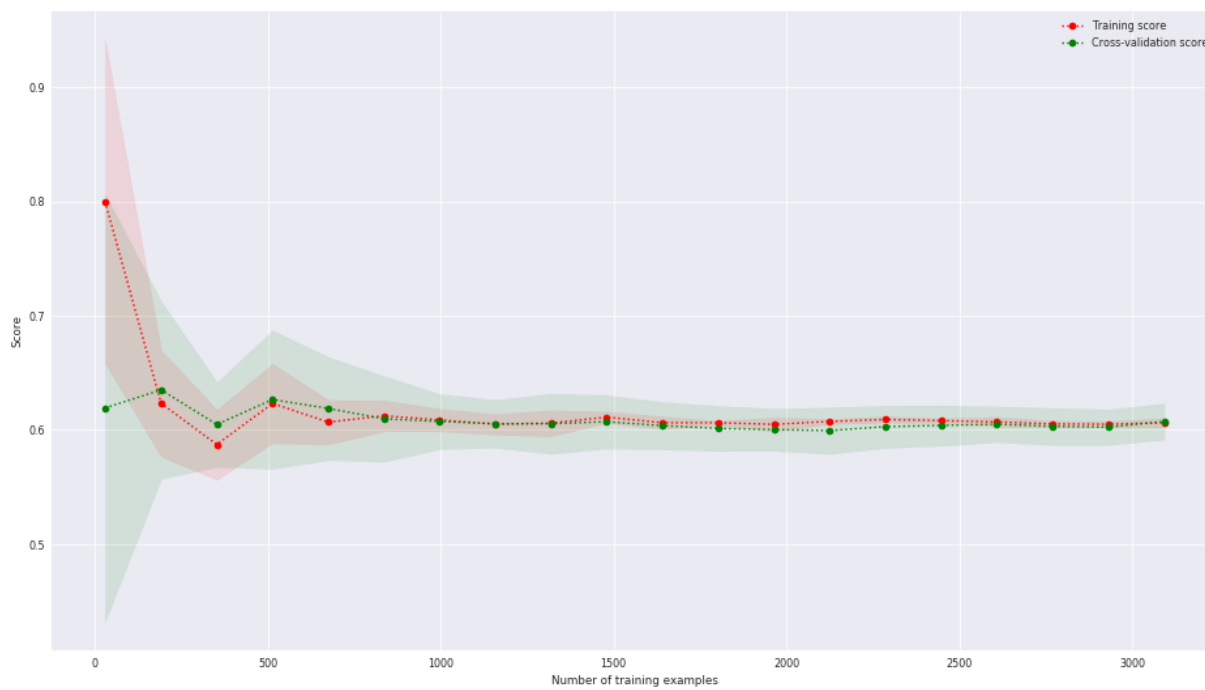


Figura A.3: EMA 89 200, QDA

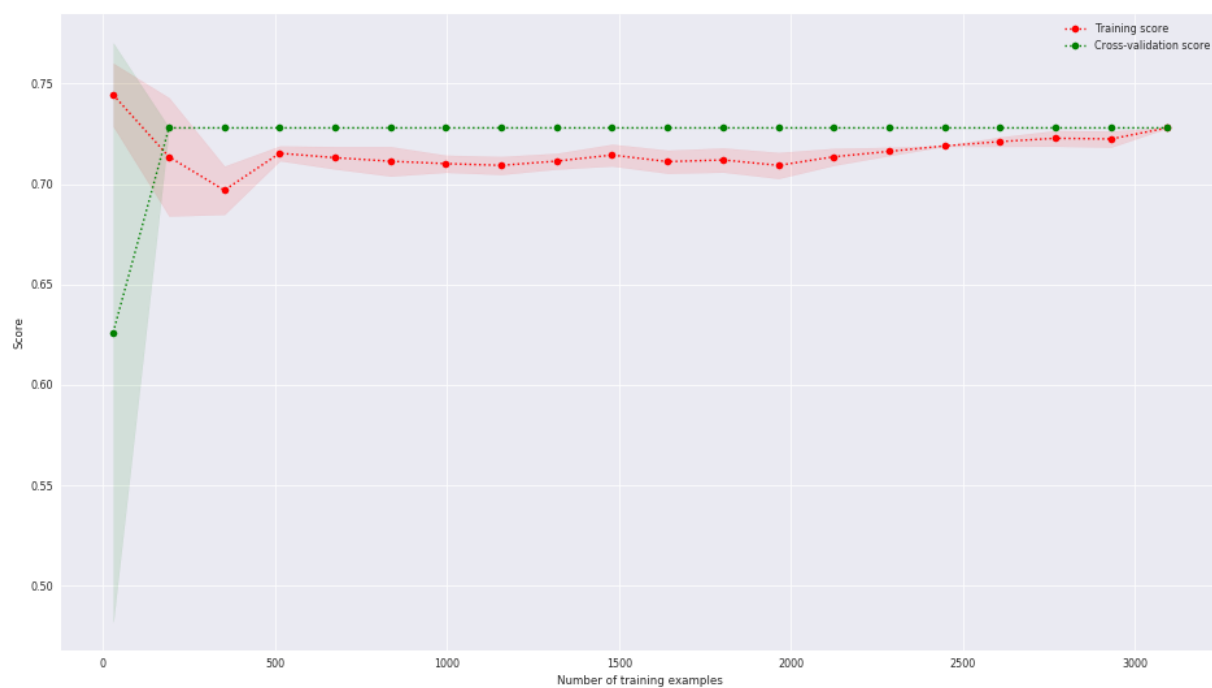


Figura A.4: EMA 89 200, SVM

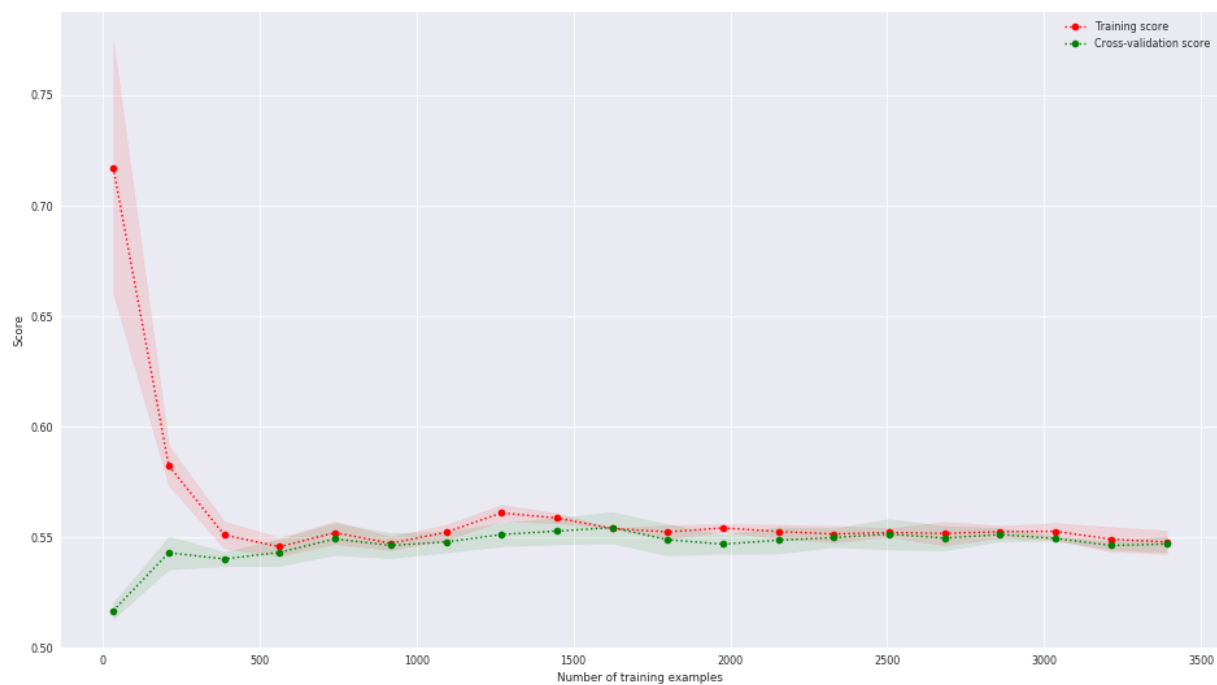


Figura A.5: MACD, Logistic Regression

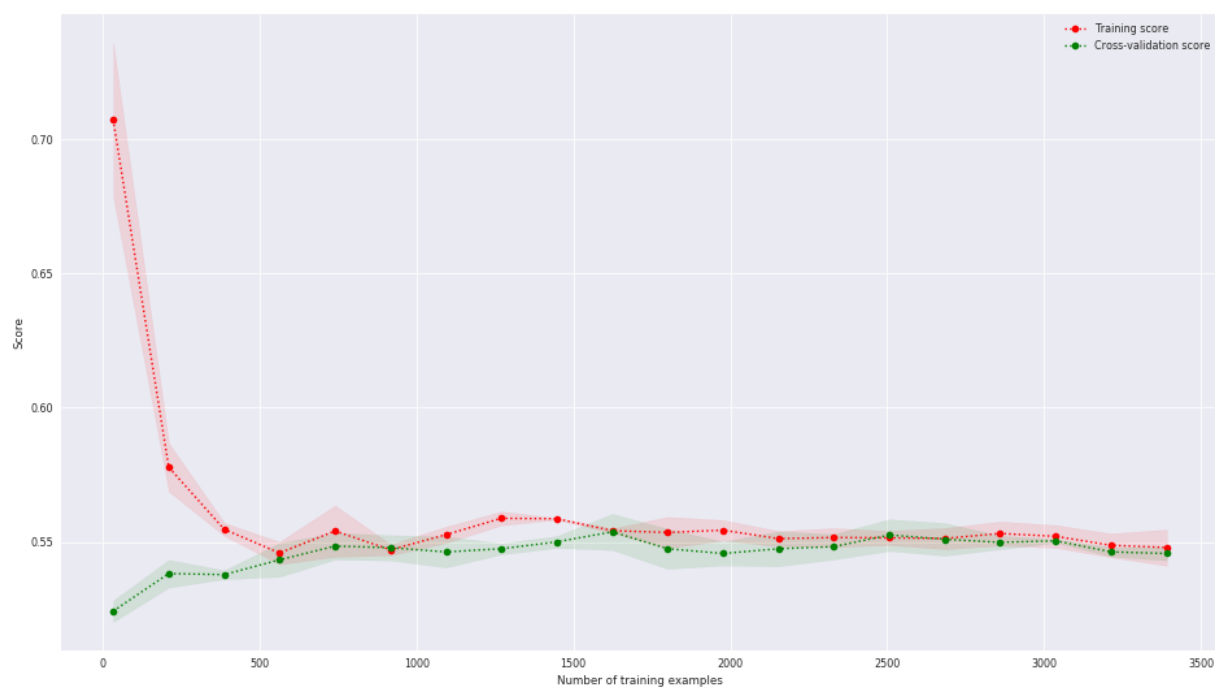


Figura A.6: MACD, LDA

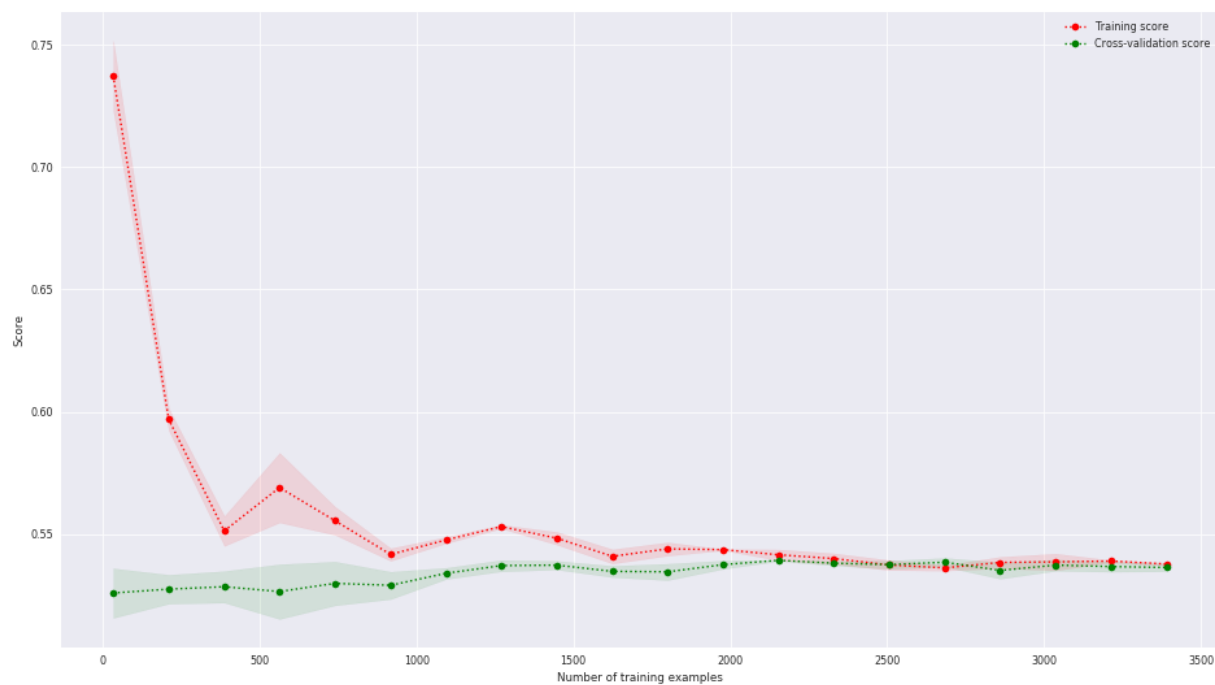


Figura A.7: MACD, QDA



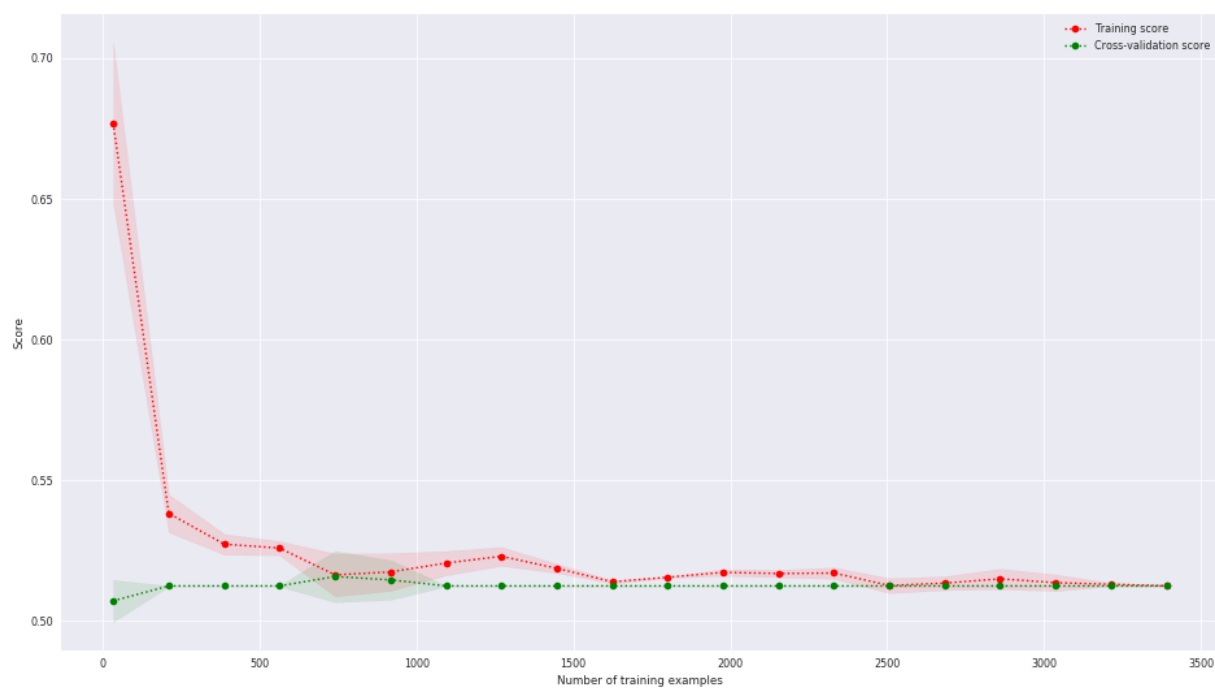


Figura A.8: MACD, SVM

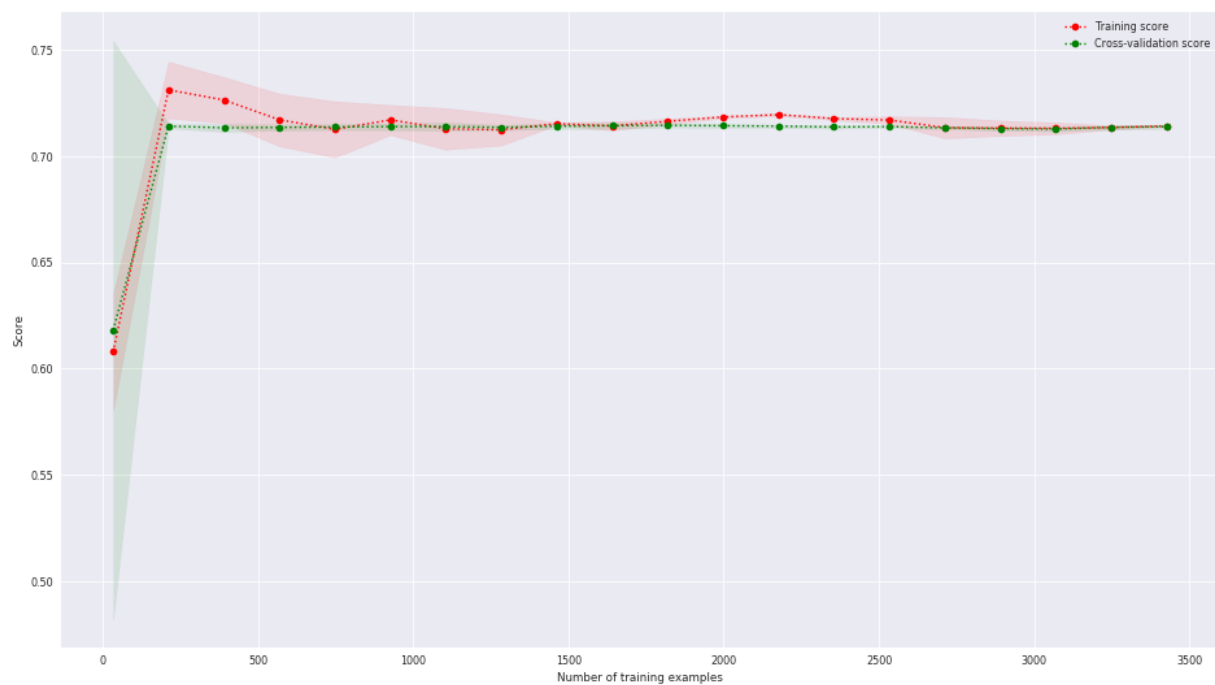


Figura A.9: RSI, Logistic Regression

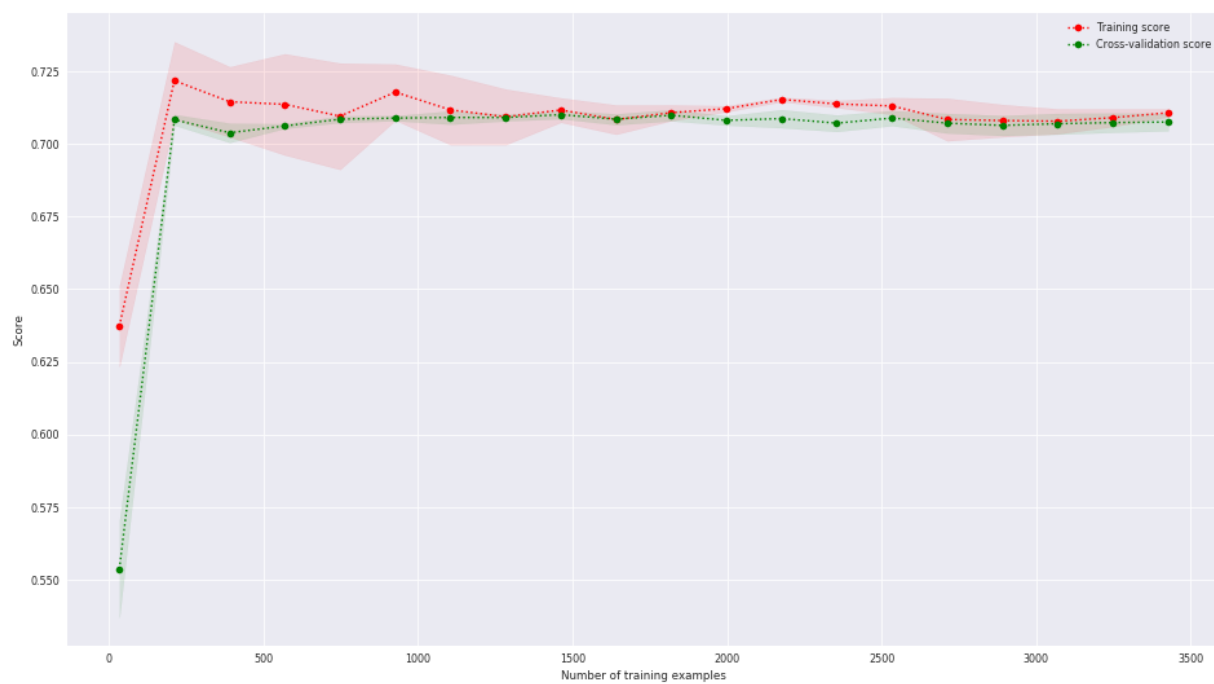


Figura A.10: RSI, LDA



Figura A.11: RSI, QDA

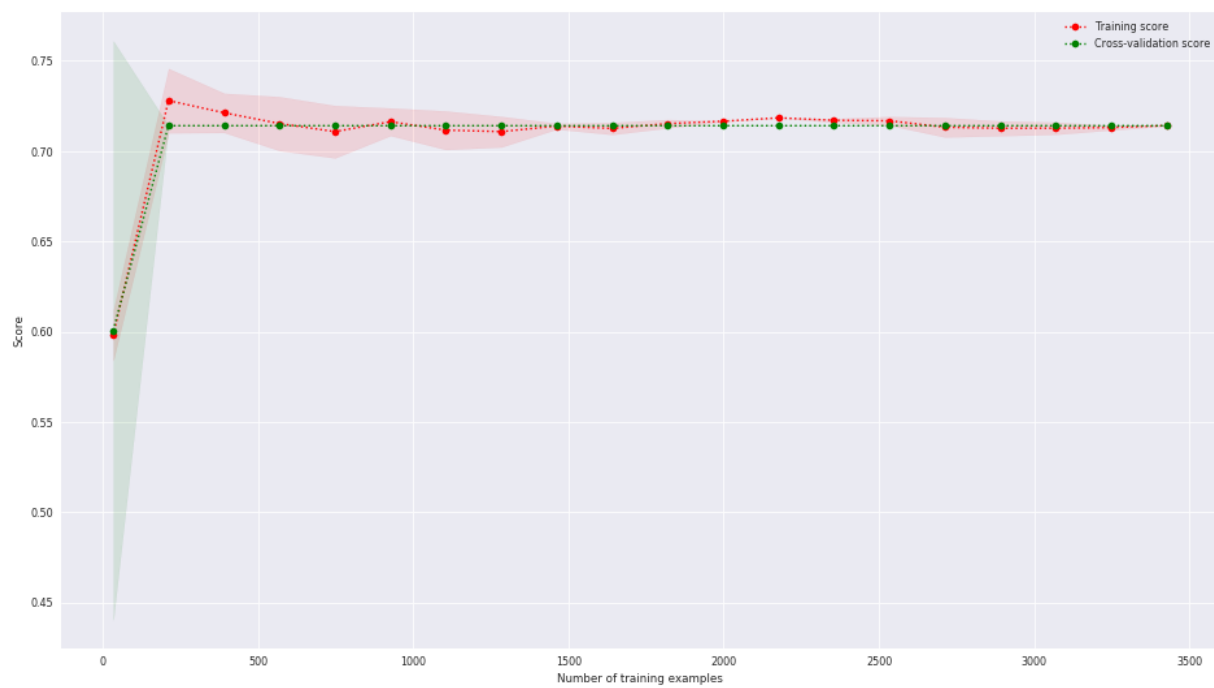


Figura A.12: RSI, SVM

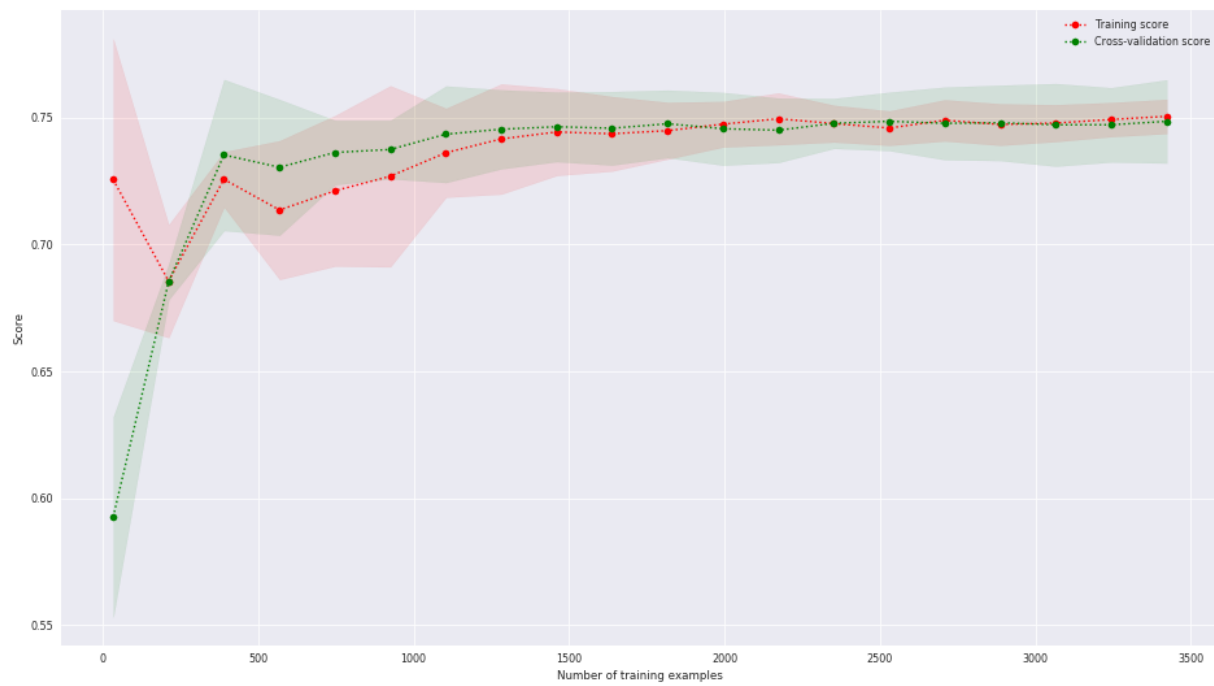


Figura A.13: SO, Linear Regression

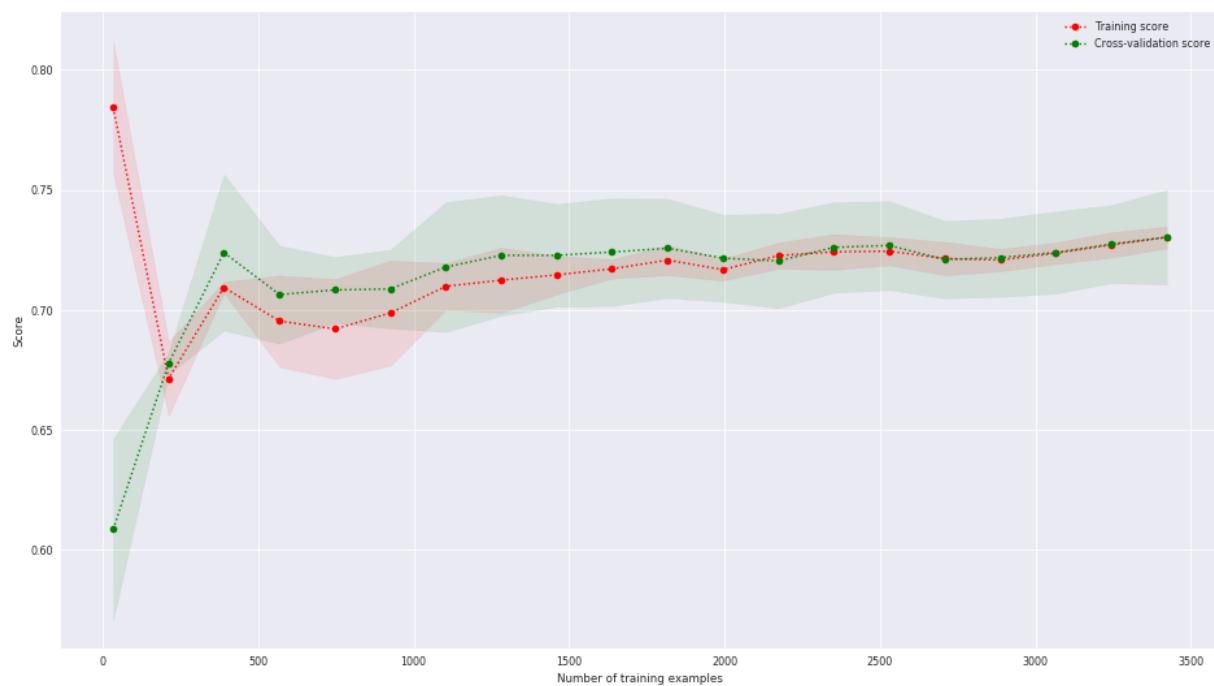


Figura A.14: SO, LDA



Figura A.15: SO, QDA

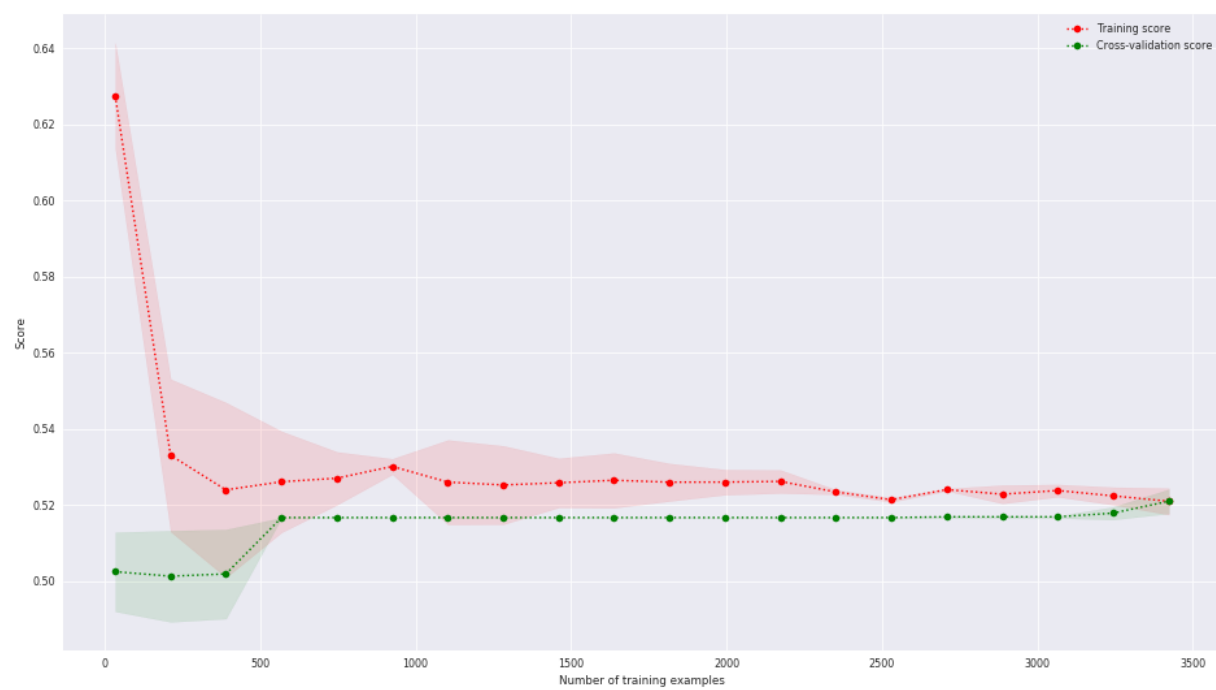


Figura A.16: SO, SVM



# Apéndice B

## Parámetros en Funciones Sklearn

En esta sección se muestran algunos de los parámetros de las funciones implementadas, con su respectiva documentación [7].

### B.1. Logistic Regression

- *penalty* : default 'l2'. Usado para especificar la norma usada en la penalización.
- *tol* : default 1e-4. Tolerancia para el criterio de parada.
- *C* : default 1.0 Inverse of regularization strength.
- *max\_iter* : default 100. Máximo número de iteraciones tomados por los *solvers* para converger.

### B.2. Linear Discriminant Analysis

- *tol* : default 1.0e-4. Umbral utilizado para la estimación de rango.

### B.3. Quadratic Discriminant Analysis

- *tol* : default 1.0e-4. Umbral utilizado para la estimación de rango.

### B.4. SVM

- *kernel* : default 'rbf'. Specifies the kernel type to be used in the algorithm.
- *gamma* : default 'auto'. Coeficiente de regularización del Kernel.
- *tol* : default 1e-3. Tolerancia para el criterio de parada.





# Bibliografía

- [1] *BBVA Trader*. Disponible en <https://www.bbva.com/es/invertir-bolsa-desde-cero-analisis-fundamental-analisis-tecnico/>.
- [2] *Historia Bolsa de Valores*. Disponible en <http://www.uv.mx/personal/mvalle/files/2011/08/BOLSA-DE-VALORES-E-INDICES-BURSATILES.pdf>.
- [3] *Historia Índice Bursátil*. Disponible en <http://www.encyclopediafinanciera.com/inversion/rentavARIABLE/indicesbursatiles.htm>.
- [4] *Inteligencia Artificial y Conciencia*. Disponible en [http://www3.uah.es/benito\\_fraile/ponencias/inteligencia-artificial.pdf](http://www3.uah.es/benito_fraile/ponencias/inteligencia-artificial.pdf).
- [5] *Iris*. Disponible en <https://archive.ics.uci.edu/ml/datasets/iris>.
- [6] *Simulador BVC*. Disponible en [http://www.bvc.com.co/pps/tibco/portalbvc/Home/AcercaBVC/Comunicados\\_Prensa?com.tibco.ps.pagesvc.action=updateRenderState&rp.currentDocumentID=2bb648\\_143bae672ca\\_-7c240a0a600b&rp.revisionNumber=1&rp.attachmentPropertyName=Attachment&com.tibco.ps.pagesvc.targetPage=1f9a1c33\\_132040fa022\\_-78750a0a600b&com.tibco.ps.pagesvc.mode=resource&rp.redirectPage=1f9a1c33\\_132040fa022\\_-787e0a0a600b](http://www.bvc.com.co/pps/tibco/portalbvc/Home/AcercaBVC/Comunicados_Prensa?com.tibco.ps.pagesvc.action=updateRenderState&rp.currentDocumentID=2bb648_143bae672ca_-7c240a0a600b&rp.revisionNumber=1&rp.attachmentPropertyName=Attachment&com.tibco.ps.pagesvc.targetPage=1f9a1c33_132040fa022_-78750a0a600b&com.tibco.ps.pagesvc.mode=resource&rp.redirectPage=1f9a1c33_132040fa022_-787e0a0a600b).
- [7] *Sklearn*. Disponible en <http://scikit-learn.org/>.
- [8] Bulla, L., (1998), 27.
- [9] Cedervall, Fredrik, (2012).
- [10] Mudry, Albert and Tjellström, Anders, *Advances in Oto-Rhino-Laryngology* **71** (2011), 1.
- [11] Nayak, S C and Misra, B B and Behera, H S, *International Journal of Computer Information Systems and Industrial Management Applications* **6** (2014), 257.
- [12] Soni, Sneha and Shrivastava, S, *International Journal on Computer ...* **02** (2010), 2942.

- [13] Tharwat, Alaa and Gaber, Tarek and Ibrahim, Abdelhameed and Hassanien, Aboul Ella, *AI Communications* **30** (2017), 169.

# Participantes

Ramiro Andrés Barrios Valencia	MSc. Ingeniería de Sistemas y Computación
Alejandro Esteban Rendón Diosa	Candidato: Ing. Sistemas y Computación